DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

DSTO

Combinatorics of Crossing Networks

Geoff A. Latham

DSTO-RR-0176

Combinatorics of Crossing Networks

Geoff A. Latham

Communications Division
Electronics and Surveillance Research Laboratory

DSTO–RR–0176

ABSTRACT

In scenarios such as software agents sharing time slots or blocks in different physical CPUs or memories, communications agents accessing channels or bandwidth on shared links or bearers and economic agents trading in commodities in financial markets, the agents in question can exchange resources among themselves for mutual benefit. This report develops a combinatorial probability model to describe the like-for-like exchange of resources between multiple unbiased agents. The expected exchange rates are computed for individual agents, syndicates of agents and the collection of all agents. These performance benchmarks provide intuition and understanding for the performance of real crossing networks. A number of combinatorial identities are also produced as a consequence of the modelling and analysis.

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury, South Australia, Australia 5108

Telephone: (08) 8259 5555
Facsimile: (08) 8259 6567

AR No. AR-011-477
June 2000

APPROVED FOR PUBLIC RELEASE

Combinatorics of Crossing Networks

EXECUTIVE SUMMARY

In Defence and commercial networks there is a trend toward user agents sharing non-exclusive access to common distributed resources (e.g. servers, databases, links etc). The modelling of the exchange of such resources between agents is therefore a problem of interest.

This report develops a new combinatorial probability model for the exchange of resources between multiple unbiased agents (or users) within an arena which facilitates such exchange (a crossing network). Once formulated, the model allows the calculation of the expected exchange rate (the mean number of exchanged resources) when an arbitrary number of agents meet. For only a few (two or three) agents, many more probabilistic properties of the exchange rate are derived.

The model uses the discrete probability theory of W. E. Deming (of post-WWII TQM fame in Japan) and G. J. Glasser from the 1950s to build a model which both further develops the theory and gives relevant insights into modern electronic exchange applications.

The expected exchange rate plots which result provide benchmark performance curves, indexed by the parameters of the model, for applications such as concurrent processes sharing memory or CPU resources, communications agents sharing channel or band-width resources on common links or bearers and economic agents with a shared interest in trading financial commodities.

Being a mathematical treatment, the report is necessarily, compact, brief and concise. Although motivated by physical applications (as mentioned above), these are better discussed in separate forums. The implications for these applications are, however, real, tangible and of significance to users of the corresponding applied systems.

Authors

Geoff A. Latham
Communications Division

The author joined the Communications Division of DSTO in 1996 subsequent to working for a number of years in industry on quantitative models and a period of postdoctoral work. He was a member of the Information Access Group, Secure Communications Branch, and, after managing a task on multi-channel digital technology and array processing algorithms for HF and other communications bands, moved to the Signals Analysis Group within the same branch in 1999.

His current research interests include applicable mathematics and statistics in data manipulation, information theory, inverse problems, computational biology and capital markets.

Contents

Glossary **xi**

1 Introduction **1**

1.1 Notation and definitions 2

1.2 Assumptions and problem statement 3

2 Preliminary results **5**

2.1 A Probability Lemma 5

2.2 The Deming-Glasser distribution 7

2.3 A special exchange problem 8

3 A few agents **11**

3.1 The 2 agent case 11

3.1.1 The exchange distribution and its moments 11

3.1.2 Small mean behaviour 13

3.1.3 Estimation of common names 14

3.2 The 3 agent case 14

3.2.1 Common name distributions 14

3.2.2 Combinatorial identities 16

3.2.3 The exchange problem 17

3.2.4 Two special cases 19

4 The general case **20**

4.1 All agents 20

4.2 A single agent 22

4.3 A syndicate of agents 24

4.4 The number of common names 26

4.5 Reductions under EDCI 28

5 Conclusion **32**

References **33**

Appendices

A	**Some Computational Details**	**34**
A.1	General identities	34
	A.1.1 Recursive evaluation of (A1)	37
A.2	Evaluation of $T_{lm}(r)$	39
	A.2.1 Method 1	39
	A.2.2 Method 2	40
B	**Derivations and Identities**	**41**
B.1	Derivation of $P(c_{lm}, c_{lq})$ for $L = 3$	41
B.2	Elementary identities	42
B.3	Alternative derivation for all agents	43

Figures

1	A schematic for $t = 3$ showing three samples of sizes n_l taken from overlapping pools of size N_l - dark shading is the random variable c while light shading indicates C.	9
2	For the parameter choices $m = n$, $N = M = 100$ and $\mathcal{C} = 50$, (a) plots the distributions $\{p_j\}$ for the three values $n = 30$, 50, 100 (left to right) while (b) plots the moments μ, σ^2 and γ as functions of n.	13
3	For the values $N = 100$ and $\mathcal{C} = 50$, (a) plots the three means $\mu/\mathcal{C}$, $\mu_\ell/\mathcal{C}$ and $\frac{1}{2}\lambda^2$ ($L = 2$) (top to bottom) while (b) plots the variances $\sigma^2/\mathcal{C}$ and $\sigma_\ell^2/\mathcal{C}$ as the curves exhibiting maxima ($\sigma^2 > \sigma_\ell^2$) and the variance from (3) for $L = 2$, all as functions of the state list fraction $\lambda \equiv n/N$ (see the text).	20
4	(a) A plot of $\mu/\mathcal{C}$ for $L = 2, \ldots, 10$ (bottom to top); (b) plot of the relative margin $\Delta_L\mu(L-1)/\mu(L-1)$ versus $\rho\lambda$ for $L = 3, \ldots, 11$ (top to bottom).	30
5	(a) A plot of the efficiency $e(\rho\lambda)$ for $L = 3, \ldots, 10$ (bottom to top); (b) efficiency contours in L-$\rho\lambda$ space for $\epsilon = 0.95, 0.90, \ldots, 0.55$ (top to bottom).	30
6	(a) A plot of $\mu_\ell/(\rho\lambda_\ell\mathcal{C})$ (bottom to top); (b) the single agent relative margin $\Delta_L\mu_\ell(L-1)/\mu_\ell(L-1)$ as functions of $\rho\lambda$ for $L = 3, \ldots, 10$ (top to bottom).	31
7	A plot of the efficiency difference $e_\ell - e$ as a function of $\rho\lambda$ for the values $L = 3, \ldots, 10$.	31

8 Plots of the means μ_{I_s} in (a) and the efficiencies e_{I_s} in (b) as functions of $\rho\lambda$ for the values $s = 1, \ldots, 10$ with $L = 10$. 32

Glossary

LTP Law of Total Probability ([4])

CL Combinatorial Lemma (section B.2)

PL Probability Lemma (section 2.1)

EDCI Exponential Decay of Common Interest (section 1.1)

1 Introduction

A crossing network is the generic name for a venue or mechanism (real or virtual) where the current 'order and disposal' states of agents are compared and resources exchanged - resources are 'crossed' between agents as compatibilities in required and released resources are found. Such a situation is typical of exchange between software agents or exchange in physical markets.

A general statement of the *exchange problem* is as follows. Consider $L \geq 2$ agents $a_1, \dots, a_L$ who use resources only from their own respective pools $P_1, \dots, P_L$. Each pool P_l consists of a collection $\vartheta_1, \dots, \vartheta_{N_l}$ of resource names. Each of these resources is *divisible* in the sense that multiple agents $a_1, \dots, a_t$ may simultaneously use a different piece of a resource ϑ_i if its name appears in the pools $P_1, \dots, P_t$. With each agent a_l, there is associated a current state list: $(s_1\vartheta_1, \dots, s_{n_l}\vartheta_{n_l})$, $s_i \in \{+1, -1\}$, of length n_l with every ϑ_i belonging to P_l and occurring at most once. If $s_i = +1$, agent a_l *requires* resource ϑ_i, while if $s_i = -1$, agent a_l releases resource ϑ_i. The state list is therefore an 'order and disposal' list representing the current state of required and surplus resources for agent a_l. Such a list is used by any resource controller/allocator serving the agents. A resource name ϑ_i is *exchangeable* between agents a_l and a_m if ϑ_i is in both state lists and s_i in a_l's list is of opposing sign to s_i in a_m's list, i.e. ϑ_i is simultaneously surplus to and required by different agents. Any system of agents would seek maximum exchangeability to maximise utilisation of the resources available through exchange between agents and thereby avoid having to either create unavailable resources or waste unrequired resources on offer in the state lists.

Consider the simultaneous comparison between all agents' state lists. The *exchange problem* (or *crossing problem*) is one of performance evaluation which asks for the *exchange rates* given by:

(a) the expected number of resource names exchanged between all L agents;

(b) the expected number of resource names exchanged within the state list of a specific agent a_ℓ; and, more generally,

(c) the expected number of resource names exchanged within the state lists of the specific agents $a_{\ell_1}, \dots, a_{\ell_s}$, $2 \leq s \leq L-1$.

Determining these rates requires the calculation of probabilistic expectations as performance benchmarks of crossing networks. Clearly the result is greatly influenced by the prior distribution assigned to the occurrence of resource names in state lists. The analysis here makes the *unbiased* assumption of equal likelihood of all resource names and equal likelihood of the require/release intention for each resource. This 'random' collection of agents might be called typical of non-cooperative (independent) agents and so provides an unbiased benchmark for the crossing network's performance.

This report gives an account of the mathematical detail of the combinatorial probability model of the exchange of like-for-like resources among many unbiased agents. While the crossing problem considered here is motivated by a concrete real-world application, the formulated mathematical and probability model, although an abstraction, provides

concrete exact results from which intuition on the behaviour of real crossing networks can be gained.

Under the unbiased assumptions, some analytical results are available for general L. For $L = 2$ and 3 the means in (a) and (b) are computed along with some higher moments. For arbitrary L, only the means are given and it remains open to characterise the higher moments of the random variable describing the number of exchangeable resource names. Even so, the derivation of the means and some higher moments produce some interesting combinatorial identities along the way.

Some specific applications of the crossing network scenario are as follows:

- Each agent a_l is a concurrent software process which may request/release time slots on different CPUs, memory blocks of different physical memories, or record blocks in different databases (ϑ_i);
- Each agent a_l on a communications network may require/release dedicated channels or capacity on different physical or logical links or bearers (ϑ_i);
- In market exchange, if price discovery is efficient, the agents a_l may exchange (trade) like commodities (ϑ_i) among themselves.

1.1 Notation and definitions

Let $i_1 \cdots i_t$, $2 \le t \le L$, be the multi-index designator for the (distinct) agents $a_{i_1}, \dots, a_{i_t}$. A multi-index, being a combination from a subset of agents, never has any two entries equal. Summations with multi-index subscripts are therefore taken over all combinations from a specified agent subset. When considering such a subset $I_s = \{i_1, \dots, i_s\}$, write a_{I_s} for the corresponding agents and subscript associated quantities with I_s. If $s = 1$, denote $I_s = \{\ell\}$ by simply ℓ, while if $s = L$ (all agents) we drop the subscript entirely. For I_s fixed, the agents a_{I_s} are called a *syndicate* when $2 \le s < L$.

Let $C_{i_1 \cdots i_t}$ be the number of resource names common to each of the pools $P_{i_1}, \dots, P_{i_t}$. To assist in visualising results, it is sometimes useful to collapse these many parameters through the following functional dependence.

Definition. The L agents $a_1, \dots, a_L$ are said to have *exponential decay of common interest* (EDCI) if $C_{i_1 \cdots i_t} = \rho^t C$, for all $2 \le t \le L$ with C and $0 \le \rho \le 1$ fixed constants, for every multi-index $i_1 \cdots i_t$.

If $0 < \rho < 1$, EDCI conveniently diminishes the size of pool overlaps with an increasing number of agents considered.

Let $\lambda_l := n_l/N_l$ be the fraction of pool P_l represented in the state list of agent a_l. A collection of L agents is described by $\sum_2^L \binom{L}{t}$ overlap parameters ($C_{i_1 \cdots i_t}$), L list sizes (n_l) and L pool sizes (N_l) for a total of $2^L + L - 1$ parameters. Define also the quantity $\pi_k(i_1, \dots, i_t) := \prod_{r=1}^t (n_{i_r} - k)/(N_{i_r} - k)$ so that, for example, $\pi_0(i_1, \dots, i_k) = \lambda_{i_1} \cdots \lambda_{i_k}$.

Definition. The resource name ϑ_i is said to be *uniquely common* to the state lists of agents $a_{i_1}, \dots, a_{i_t}$ if it appears in the state lists of these agents and no others.

Denote by $c_{i_1 \cdots i_t}$ the number of names common to the state lists of agents $a_{i_1}, \dots, a_{i_t}$ and by $d_{i_1 \cdots i_k}$ the number of names uniquely common to the state lists of agents $a_{i_1}, \dots, a_{i_t}$. The utility of the concept of unique commonality is that, since the partition of common names among state lists into the uniquely common names among state lists is a disjoint one, cumulative counting of uniquely common names avoids double counting when determining the number of names common to two or more agents within a specified subset of all agents.

Let J (respectively, J_ℓ, J_{I_s}) be the random variable of the number of exchangeable names within the state lists of all L agents (respectively, agent a_ℓ, the syndicate a_{I_s}) and let the mean be $\mu = E[J]$ (respectively, $\mu_\ell = E[J_\ell]$, $\mu_{I_s} = E[J_{I_s}]$). Let x (respectively, x_ℓ, x_{I_s}) be the random variable of the number of names within the state lists of all L agents (respectively, agent a_ℓ, the syndicate a_{I_s}) which are common to two or more state lists and let its mean be $\chi = E[x]$ (respectively, $\chi_\ell = E[x_\ell]$, $\chi_{I_s} = E[x_{I_s}]$).

Definition. The *efficiency* of the crossing network (respectively, agent a_ℓ, syndicate a_{I_s}) is $e := \mu/\chi$ (respectively, $e_\ell := \mu_\ell/\chi_\ell$, $e_{I_s} := \mu_{I_s}/\chi_{I_s}$).

The efficiency, which measures the ratio of the expected number of exchangeable names to the expected number of common names, gives expected performance as a fraction of the best expected performance possible.

Let $[c]_k := c(c-1)\cdots(c-k+1)$ be the descending factorial and $(c)_k := c(c+1)\cdots(c+k-1)$ the ascending factorial. The factorial moments of a random variable V are denoted by $\mu_{[k]} := E[[V]_k]$. Let $\left[{k \atop l}\right]$ and $\left\{{k \atop l}\right\}$ denote, respectively, the Stirling numbers of the first and second kind [7]. These are defined by the relations

$$[V]_k = \sum_{l=0}^{k} \begin{bmatrix} k \\ l \end{bmatrix} V^l \qquad \text{and} \qquad V^k = \sum_{l=0}^{k} \begin{Bmatrix} k \\ l \end{Bmatrix} [V]_l$$

with, in particular, $\left\{{k \atop l}\right\} := \sum_i (-1)^i \binom{l}{i} (l-i)^k / l!$. Finally, let Δ_r be the forward difference operator in r, i.e. $\Delta_r f_r = f_{r+1} - f_r$.

Remark. Note our use of ℓ versus l. When referring to a fixed and specified agent, ℓ is used as a designator, but l is the generic subscript. Hence ℓ should be interpreted as "the specific fixed agent a_ℓ".

The following acronyms are used throughout:

LTP – Law of Total Probability ([4]);

CL – Combinatorial Lemma (section B.2);

PL – Probability Lemma (section 2.1);

EDCI – Exponential Decay of Common Interest (section 1.1).

1.2 Assumptions and problem statement

The derivation of the crossing network exchange model is based upon the *unbiased* assumption which holds throughout this paper and is comprised of the following two parts:

(A1) For every agent a_l, each resource name ϑ_i in the pool P_l is equally likely to appear in their state list;

(A2) For every resource name, $\Pr(s_i = +1) = \Pr(s_i = -1) = \frac{1}{2}$.

The *unbiased exchange problem* is to find the distribution of J (respectively, J_ℓ, J_{I_s}) under the assumptions (A1) and (A2). This question is not simply answered for arbitrary L so the focus here is on the simpler question of finding the means, and where possible, the higher moments. Hence, the following questions are posed and answered:

(a) find $\mu := E[J]$;

(b) find $\mu_\ell := E[J_\ell]$; and

(c) find $\mu_{I_s} := E[J_{I_s}]$.

These expected exchange rates measure respectively, the 'system performance' which is of interest for network management, the individual agent's performance and the performance of an arbitrary subset of the agents (a syndicate). They establish expected unbiased benchmark performances for systems in which like-for-like exchange takes place. In particular, for applications where exchange is promoted or controlled by system processes, this benchmark represents the 'random outcome' performance level which should always be exceeded. If such performance is not reached, both agents and system management or regulators would have cause for concern.

2 Preliminary results

This section gives some elementary discrete probability results which follow from the unbiased assumptions above. The first result, called the Probability Lemma (PL), is pivotal for all that follows in that it facilitates the use of the results of Deming and Glasser [1]. These results are surveyed in section 2.2.

2.1 A Probability Lemma

The equi-probable assumption (A2) implies a useful anonymity when dealing with state lists. The results here use this assumption to relate the number of exchangeable names between state lists to the number of common names among those state lists. Given (A2), the link is a simple counting exercise.

Let U_t (respectively, U_s) be the random variable of the number of names common to the t (respectively, s) given state lists of agents $a_{i_1}, \cdots, a_{i_t}$ (respectively, $a_{h_1}, \cdots, a_{h_s}$) and let J_t (respectively, J_s) be the random variable of the number of exchangeable names among these U_t (respectively, U_s) common ones. The abbreviations $P(j_t|u_t) \equiv \Pr(J_t = j_t|U_t = u_t)$, $P(u_t) \equiv \Pr(U_t = u_t)$, etc are used to simplify the notation.

Probability Lemma (PL). *Under (A2), the following hold:*

(i) If a given t state lists have u_t names in common, then

$$P(j_t|u_t) = (2^{t-1} - 1)^{j_t} 2^{-(t-1)u_t} \binom{u_t}{j_t};$$

(ii) The generating functions $J_t(z) := \sum_{j_t} P(j_t) z^{j_t}$ and $U_t(z) := \sum_{u_t} P(u_t) z^{u_t}$ are related by

$$J_t(z) = U_t(2^{1-t} + (1 - 2^{1-t})z),$$

so that

$$\begin{aligned} E[[J_t]_k] &= (1 - 2^{1-t})^k E[[U_t]_k], \quad k = 0, 1, 2, \dots, \\ E[J_t] &= (1 - 2^{1-t}) E[U_t], \\ E[J_t^2] &= (1 - 2^{1-t})^2 E[U_t^2] + 2^{1-t}(1 - 2^{1-t}) E[U_t]; \end{aligned}$$

(iii) For $i_1 \cdots i_t \neq h_1 \cdots h_s$ as combinations, the covariance of J_t and J_s satisfies

$$E[J_t J_s] = (1 - 2^{1-t})(1 - 2^{1-s}) E[U_t U_s].$$

Proof. The proofs of (ii) and (iii) rely on (i) which is proved first. Suppose that t lists have u_t names in common, then under the equi-probable assumption (A2), $P(j_t|u_t)$ is just the number of signed configurations of the t lists with u_t common names for which j_t names are exchangeable divided by the total number of possible signed configurations of the lists.

We need only consider the configurations among the common names in computing this ratio. Let $\vartheta_1, \dots, \vartheta_{u_t}$ be the common names and consider the configurations of the lists:

$$\{s_1^{(1)}\vartheta_1, \dots, s_{u_t}^{(1)}\vartheta_{u_t}\},$$
$$\vdots$$
$$\{s_1^{(t)}\vartheta_1, \dots, s_{u_t}^{(t)}\vartheta_{u_t}\}.$$

There are 2^{tu_t} such possible configurations for $s_i^{(l)} \in \{+1, -1\}$. Suppose there are j_t exchangeable names. There are exactly $\binom{u_t}{j_t}$ ways of choosing the exchangeable names each of which has 2^{j_t} possible ways of choosing the s_i's for these j_t names and $2^{u_t - j_t}$ possible ways of choosing the s_i's for the remaining $u_t - j_t$ non-exchangeable names. Hence, $2^{u_t-j_t}\binom{u_t}{j_t}2^{j_t}$ is the number of possible configurations of *one* list with j_t exchangeable names. Now fix a configuration of the first list and consider the remaining $t-1$ lists. For each of the j_t names, ϑ_i say, that are exchangeable, there are $2^{t-1} - 1$ configurations of the other $t-1$ lists in which that name has at least one list with an opposing sign $s_i^{(l)}$ in some list, i.e. 2^{t-1} possible configurations less the one configuration where ϑ_i has the same sign in all the remaining $t-1$ lists. Hence there are $(2^{t-1}-1)^{j_t}$ possible configurations in which j_t names are exchangeable with the first list fixed. The number of possible configurations with j_t exchangeable names among the u_t common names is therefore $\nu_{j_t} := (2^{t-1}-1)^{j_t}2^{u_t}\binom{u_t}{j_t}$, and since $\sum_j \nu_j = 2^{tu_t}$, then $P(j_t|u_t)$ is the ratio of these two numbers as required by (i).

To derive (ii), apply the LTP to write the generating function for J_t as

$$J_t(z) = \sum_{j_t, u_t} z^{j_t} P(j_t|u_t)P(u_t).$$

Now use the explicit form in (i) to write $\sum_{j_t} z^{j_t}P(j_t|u_t) = (2^{1-t} + (1-2^{1-t})z)^{u_t}$ so that

$$J_t(z) = \sum_{u_t} P(u_t)(2^{1-t} + (1-2^{1-t})z)^{u_t} = U_t(2^{1-t} + (1-2^{1-t})z).$$

Since $J_t(1+z)$ generates the factorial moments of J_t, and $J_t(1+z) = U_t(1 + (1-2^{1-t})z)$, then the factorial moments are related by scaling with powers of $1 - 2^{1-t}$. The cases $k=1$ and $k=2$ then imply the given relations between the first and second moments.

To prove (iii), again apply the LTP and Bayes' rule to get

$$\begin{aligned} P(j_t|j_s) &= \sum_{u_t} P(j_t|u_t)P(u_t|j_s) = \sum_{u_t,u_s} P(j_t|u_t)P(u_t|u_s)P(u_s|j_s) \\ &= \sum_{u_t,u_s} P(j_t|u_t)\frac{P(u_t,u_s)}{P(u_s)}\frac{P(j_s|u_s)P(u_s)}{P(j_s)} \end{aligned}$$

and so

$$P(j_t, j_s) = \sum_{u_t,u_s} P(j_t|u_t)P(j_s|u_s)P(u_t,u_s).$$

Now taking the required expectation gives

$$\begin{aligned} E[J_tJ_s] &= \sum_{u_t,u_s}\sum_{j_t,j_s} j_tP(j_t|u_t)j_sP(j_s|u_s)P(u_t,u_s) \\ &= (1-2^{1-t})(1-2^{1-s})\sum_{u_t,u_s} u_tu_sP(u_t,u_s), \end{aligned}$$

where again, the expression in (i) has been used to evaluate the j-sums. □

Observe that the conditional probability in (i) of the PL is in fact a Binomial distribution on u_t trials with the trial success probability $\rho = 1-2^{1-t}$ and the trial failure probability $1-\rho = 2^{1-t}$.

2.2 The Deming-Glasser distribution

Before the advent of modern computing resources for record keeping and manipulation, sampling methods to assist with counting problems were a topical issue of research. For example, the works of Deming and Glasser [1] and Goodman [2] were motivated by the problem of comparing large lists of names to determine the size of common membership through the comparison of small random samples taken from those large lists. Two particular motivating applications for this work were:

- determining the number of names common to different lists of social welfare recipients;
- determining the number of books common to different library catalogues.

In their day, these posed significant problems because of the lack of ease of comparing large collections of paper records. This section gives a brief summary of those of Deming and Glasser's results from [1] which are used below in the crossing network model.

Consider t overlapping pools of names $P_1,\dots,P_t$ of respective sizes $N_1,\dots,N_t$ with C names in common (see Figure 1). For t independent respective random samples of sizes $n_1,\dots,n_t$ taken from these pools, Deming and Glasser [1] showed that the random variable c of the number of names common to the t samples has the distribution

$$P_t(C;c) = \sum_{k=c}^{C}(-1)^{k-c}\binom{k}{c}\binom{C}{k}\prod_{l=1}^{t}\frac{\binom{n_l}{k}}{\binom{N_l}{k}}, \quad c = 0,1,\dots,\min_l(n_l,C)$$

which is called the DG-distribution. Some simple rearranging of terms shows that the generating function $P_t(z) := \sum_c P_t(C;c)z^c$ is

$$P_t(z) = {}_{t+1}F_t(-C;-n_1,\dots,-n_t;-N_1,\dots,-N_t;1-z),$$

where ${}_pF_q$ is the generalised hypergeometric function [5]. This identifies the DG-distribution as a generalised hypergeometric factorial moment distribution (GHFD) [3]. Since $P_t(1+z)$ generates the factorial moments $\mu'_{[k]} := E[[c]_k]$, then immediately,

$$\mu'_{[k]} = [C]_k\prod_{l=1}^{t}\frac{[n_l]_k}{[N_l]_k}, \quad k = 0,1,\dots,\min_l(n_l,C).$$

Alternatively, since $\mu'_{[k]} = k!E[\binom{c}{k}]$, the factorial moments can be identified by applying the Lemma of section B.2 with $b_k = \mu'_{[k]}/k!$ and $p_r = P_t(C;r)$. The derivation of $P_t(C;c)$ is a straight forward application of the LTP (cf. section 3.1). Consequently, $P_t(C;c)$ satisfies a recursion which is now discussed.

If an additional $(t+1)$-st independent random sample of size n_{t+1} is taken from a pool P_{t+1} of size N_{t+1}, still with C names in common with the pools $P_1, \ldots, P_t$, then, by the LTP,

$$P_{t+1}(C;c) = \sum_{k=0}^{C} \frac{\binom{k}{c}\binom{N_{t+1}-k}{n_{t+1}-c}}{\binom{N_{t+1}}{n_{t+1}}} P_t(C;k), \quad c = 0, 1, \ldots, \min_l(n_l, C). \tag{1}$$

Moreover, for each $r = 0, \ldots, t$, the recursion

$$P_{t+1}(C;c) = \sum_{k=0}^{C} P_{r+1}(k;c)P_{t-r}(C;k)$$

follows from the implied repeated sum representation of $P_{t+1}(C;c)$ derived by iterating (1). In this recursion, $P_0(C;k) = \delta_{kC}$ and $P_1(C;k) = \binom{C}{k}\binom{N_1-C}{n_1-k}/\binom{N_1}{n_1}$. More explicitly, augmenting the notation to include all of the omitted parameters,

$$P_p\left(\begin{matrix} C; & n_1, & \cdots, & n_p| & \\ & N_1, & \cdots, & N_p| & c \end{matrix}\right) =$$

$$\sum_k P_{p-q}\left(\begin{matrix} k; & n_{q+1}, & \cdots, & n_p| & \\ & N_{q+1}, & \cdots, & N_p| & c \end{matrix}\right) \cdot P_q\left(\begin{matrix} C; & n_1, & \cdots, & n_q| & \\ & N_1, & \cdots, & N_q| & k \end{matrix}\right)$$

$0 \leq q \leq p$, shows this last recursion in detail. The main property of the DG-distribution required here is its mean:

$$E[c] = \mu'_{[1]} = C \prod_{l=1}^{t} \frac{n_l}{N_l}. \tag{2}$$

2.3 A special exchange problem

This section considers a more restrictive version of the resource exchange problem in order to illustrate the use of the DG-distribution in obtaining explicit results for the number of exchangeable names among agents. The special problem solved here is to determine the distribution of the number of exchangeable names only within those names common to all of a given set of resource pools.

Let $\tilde{J}_t$ be the random variable of the number of exchangeable names in the state lists of agents $a_1, \ldots, a_t$, $2 \leq t \leq L$, among those C (drop the indexing notation $C_{1\cdots t}$ of section 1.1) names common to the pools $P_1, \ldots, P_t$ and let c be the random variable of the number of names common to all t state lists. Then by the LTP and the PL,

$$\begin{aligned}\Pr(\tilde{J}_t = j) &= \sum_c P(j|c)P_t(C;c) \\ &= (2^{t-1}-1)^j \sum_{c=0}^{C} \binom{c}{j} 2^{-(t-1)c} P_t(C;c),\end{aligned}$$

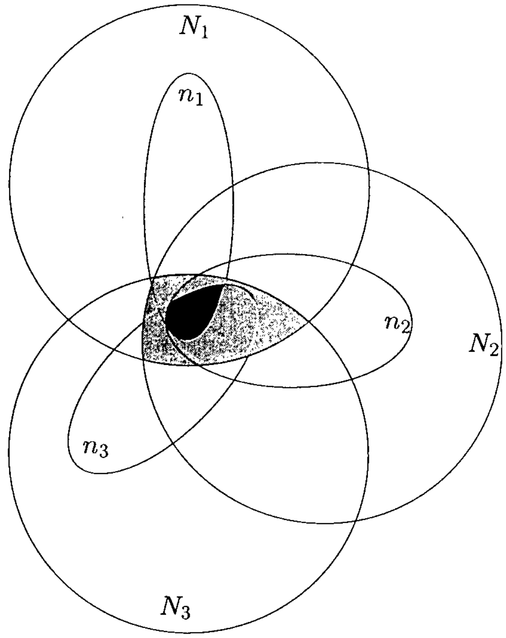


Figure 1: A schematic for $t = 3$ showing three samples of sizes n_l taken from overlapping pools of size N_l - dark shading is the random variable c while light shading indicates C.

where $P_t(C; c)$ is the DG-distribution with the parameters for agents $a_1, \ldots, a_t$. Forming the sum $\sum_j z^j \Pr(\tilde{J}_t = j)$, and evaluating the j-sum, it follows that the generating function for $\tilde{J}_t$ is

$$\tilde{J}_t(z) = {}_{t+1}F_t(-C; -n_1, \ldots, -n_t; -N_1, \ldots, -N_t; \tfrac{1}{2}(1-z))$$

and the central moments $\tilde{\mu}_k$ of $\tilde{J}_t$, derived from the factorial moment generating function $\tilde{J}_t(1+z)$ [4, 3], are

$$\tilde{\mu}_k = \sum_{i=0}^{k} (1 - 2^{1-t})^i [C]_i \prod_{l=1}^{t} \frac{[n_l]_i}{[N_l]_i} \sum_{l=0}^{k-i} \left\{ {k-l \atop i} \right\} \binom{k}{l} (-\tilde{\mu})^l, \quad k = 2, 3, \cdots,$$

with, in particular, the mean and variance being given by

$$\tilde{\mu} = (1 - 2^{1-t}) C \prod_{l=1}^{t} \frac{n_l}{N_l},$$

$$\tilde{\sigma}^2 = \tilde{\mu}_2 = \tilde{\mu}\big[1 - \tilde{\mu} + (1 - 2^{1-t})(C-1) \prod_{l=1}^{t} \frac{n_l - 1}{N_l - 1}\big].$$

Thus, the distribution and the moments of the number of exchangeable names in those names common to all pools of a given set of agents can be written down explicitly using the DG-distribution. The difficulty in determining the total number of exchangeable names overall lies in not double counting these names when they lie in multiple overlaps of the pools of many agents (cf. Figure 1).

Finally, it follows from the expressions for $\tilde{\mu}$ and $\tilde{\sigma}^2$ that for an observation j of $\tilde{J}_t$, the estimator $\hat{C} := j(1-2^{1-t})^{-1}\prod_l \lambda_l^{-1}$ is an unbiased estimator of C with standard error

$$\text{var}\,\hat{C} = \frac{C(1-2^{1-t})^{-1}}{\prod_l \lambda_l}\Big[1+(C-1)(1-2^{1-t})\prod_l \frac{n_l-1}{N_l-1}\Big] - C^2.$$

and an unbiased estimator of this standard error is:

$$\hat{V} := \hat{C}\Big[(1-2^{1-t})^{-1}\prod_l \frac{N_l-1}{n_l-1} - 1\Big] + \hat{C}^2\Big[1-\prod_l \lambda_l \frac{N_l-1}{n_l-1}\Big]$$

(cf. [4]). Clearly, the agents $a_1, \cdots, a_t$ can be replaced by any subset $a_{i_1}, \cdots, a_{i_t}$ of the L agents and the above results still hold with the replacements $\lambda_l \to \lambda_{i_l}$ and $C \to C_{i_1 \cdots i_t}$.

3 A few agents

We begin a study of the unbiased exchange problem by considering the two cases $L = 2$ and $L = 3$. Apart from introducing and illustrating the techniques used in the general case, these cases are more easily solved to reveal more moments of the exchange random variables J and J_ℓ. In particular, the basic case $L = 2$ is completely solvable through the characterisation of the exchange distribution for J, while for the $L = 3$ case, only certain common name joint distributions and the first and second moments are calculated.

3.1 The 2 agent case

The two agent case is fundamental to exchange in many agent scenarios: any exchange ultimately takes place between pairs of agents as they are compared or meet in a network. This case, therefore, forms an intuitive building block of more complicated situations. For $L = 2$, the questions (a), (b) and (c) of section 1.2 collapse to the one question since every exchangeable name in one list is in the other (i.e. $J = J_\ell$). Since there is only one overlap of resource pools, the solution of the exchange problem for this case has already been given in section 2.3 (set $t = 2$). However, a separate derivation is instructive. After computing the exchange distribution and all of its moments, the questions of small mean behaviour and the estimation of the number of names common to the two resource pools are briefly considered.

3.1.1 The exchange distribution and its moments

To aid in simplification, the notation here departs slightly from that in section 1.1. Consider two agents a_1 and a_2 with state lists of length n and m respectively consisting of respective samples of resource names from the pools P_1 and P_2 which are of respective sizes N and M. The quantities $\pi_i(1, 2)$ are written simply as π_i and $\lambda_1 = n/N$ and $\lambda_2 = m/M$.

Let j be the number of exchangeable names and suppose that P_1 and P_2 have $\mathcal{C}$ names in common. Write $p_j = \Pr(J = j)$. In this special case, a complete characterisation of p_j is readily available. In fact, we show that for two agents,

$$p_j = \sum_{k=0}^{m} 2^{-k} \binom{k}{j} \sum_{c=0}^{\mathcal{C}} \frac{\binom{c}{k}\binom{M-c}{m-k}}{\binom{M}{m}} \frac{\binom{\mathcal{C}}{c}\binom{N-\mathcal{C}}{n-c}}{\binom{N}{n}}, \quad j = 0, \ldots, \min(m, n, \mathcal{C}).$$

The proof of the formula for p_j is an application of the LTP. Define the events:

B_c = n-list contains c of the $\mathcal{C}$ common names, $0 \leq c \leq \mathcal{C}$;

A_k = m-list has k common names with the n-list, $0 \leq k \leq m$;

D_j = j exchanges occur between the n-list and m-list, $0 \leq j \leq m$.

Then using

$$\Pr(A_k) = \sum_{c=0}^{\mathcal{C}} \Pr(A_k|B_c)\Pr(B_c),$$
$$\Pr(D_j) = \sum_{k=0}^{m} \Pr(D_j|A_k)\Pr(A_k),$$

with $\Pr(B_c) = \binom{\mathcal{C}}{c}\binom{N-\mathcal{C}}{n-c}/\binom{N}{n}$, $\Pr(A_k|B_c) = \binom{c}{k}\binom{M-c}{m-k}/\binom{M}{m}$ and $\Pr(D_j|A_k) = 2^{-k}\binom{k}{j}$ (k successful probability $\frac{1}{2}$ Bernoulli trials, cf. $t = 2$ in the PL), gives the expression for p_j. The hypergeometric probabilities appear as a reflection of the (equally likely) selection without replacement of resource names from the pool. The distribution $P_2(\mathcal{C}; k) := \Pr(A_k)$, which gives the number of common names between the two lists, is just the DG-distribution for two lists (see section 2.2).

The distribution $\{p_j\}$ is a generalised hypergeometric factorial moment distribution (GHFD) [3] and is a (probability $\frac{1}{2}$) Bernoulli generalisation of the common name distribution $P_2(\mathcal{C}; k)$ in [1]. Note the symmetry under simultaneous interchange of m with n and M with N. The distribution's properties are summarised through considering the generating function $P(z) := \sum_j p_j z^j$. Let $\mu_k := E[(J-\mu)^k]$ be the central moments with $\mu = E[J]$ the mean and σ^2 the variance. The generating function is

$$P(z) = {}_3F_2(-\mathcal{C}, -n, -m; -N, -M; \tfrac{1}{2}(1-z)),$$

and the central moments μ_k are given by

$$\mu_k = \sum_{i=0}^{k} 2^{-i} \frac{[m]_i[n]_i[\mathcal{C}]_i}{[M]_i[N]_i} \sum_{l=0}^{k-i} \binom{k}{l} \left\{ {k-l \atop i} \right\} (-\mu)^l, \quad k = 2, 3, \ldots .$$

In particular, the mean (μ), variance (σ^2) and (Fisher) skewness ($\gamma := \mu_3/\sigma^3$) are given by

$$\begin{aligned}
\mu &= \tfrac{1}{2}\lambda_1\lambda_2\mathcal{C}, \\
\sigma^2 &= \mu(1 - \mu + \tfrac{1}{2}(\mathcal{C}-1)\pi_1), \\
\gamma &= \frac{1}{\sqrt{\mu}} \frac{[1 - 3\mu + 2\mu^2 + \frac{3}{2}(1-\mu)(\mathcal{C}-1)\pi_1 + \frac{1}{4}(\mathcal{C}-1)(\mathcal{C}-2)\pi_2]}{[1 - \mu + \frac{1}{2}(\mathcal{C}-1)\pi_1]^{3/2}}.
\end{aligned} \tag{3}$$

Although seemingly complicated, these expressions for the moments follow easily from the generating function. Since $P(1+z)$ generates the factorial moments, then identifying the coefficients of z^k in $P(1+z)$ shows that $\mu_{[k]} := E[[J]_k] = 2^{-k}[m]_k[n]_k[\mathcal{C}]_k/([M]_k[N]_k)$ from which standard moment relationships [3, 4] give the expressions for the central moments μ_k.

Observe that all the results here display the symmetry identified earlier expressing the freedom to re-label the agents. The distribution $\{p_j\}$ also has a number of reductions which can be identified from the corresponding $P(z)$ - these are:

- If $\mathcal{C} = M$, then $P(z) = {}_2F_1(-n, -m; -N; \frac{1}{2}(1-z))$ which characterises the hypergeometric distribution (with parameters n, m and N) generalised by the (probability $\frac{1}{2}$) Bernoulli distribution (similarly for $\mathcal{C} = N$);

- If $n = N$, then $P(z) = {}_2F_1(-C, -m; -M; \frac{1}{2}(1-z))$ which is the hypergeometric distribution (with parameters C, m and M) generalised by the (probability $\frac{1}{2}$) Bernoulli distribution (similarly for $m = M$);
- If $n = N$ and $m = M$, then $P(z) = {}_1F_0(-C; \frac{1}{2}(1-z))$ which is a binomial distribution on C trials with success probability $\frac{1}{2}$;
- If $m = M$ and $C = N$, then $P(z) = {}_1F_0(-n; \frac{1}{2}(1-z))$ which is a binomial distribution on n trials with success probability $\frac{1}{2}$ (similarly for $n = N$ and $C = M$).

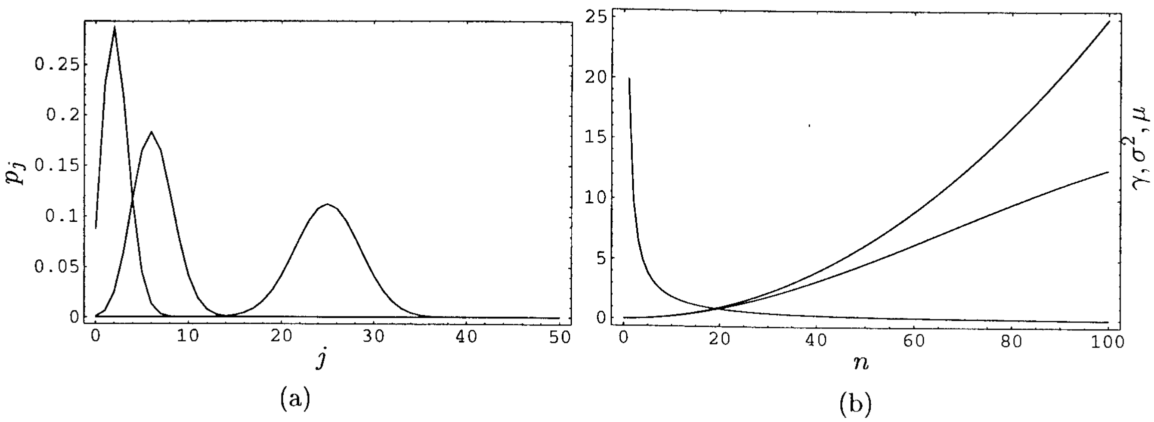


Figure 2: For the parameter choices $m = n$, $N = M = 100$ and $C = 50$, (a) plots the distributions $\{p_j\}$ for the three values $n = 30, 50, 100$ (left to right) while (b) plots the moments μ, σ^2 and γ as functions of n.

Figure 2 shows example distributions $\{p_j\}$ in (a) and the first three moments in (b). By fixing all but one of the size parameters (n), one-dimensional plots result. In these examples, the two agents are similar in that $m = n$ and $M = N$. Figure 2(a) illustrates the transition to the binomial distribution as n increases, while (b) shows the quadratic (in n) mean, an inflection in the variance and the rapidly decaying skewness.

3.1.2 Small mean behaviour

The mean and variance approximately agree for small μ. This is seen by writing

$$\sigma^2 = \mu + \mu^2 \Big[\frac{(1-m^{-1})(1-n^{-1})}{(1-M^{-1})(1-N^{-1})} (1 - C^{-1}) - 1 \Big],$$

which also illustrates that, for μ small (i.e. if any of n/N, m/M or C are small), the level curves $mn = $ const of μ also approximately preserve the variance. The degree of agreement of the two curves is shown in Figure 2(b).

The exchange distribution exhibits a favourable positive skewness which increases for decreasing mean μ (see Figure 2(b)). For the skewness,

$$\gamma = \frac{1}{\sqrt{\mu}} [1 + O(\mu)], \quad \text{as} \quad \mu \to 0.$$

Hence, the exchange distribution $\{p_j\}$ is highly positively skewed only for small mean and variance. As μ increases, γ decreases with in fact $\gamma = 0$, when, according to the reductions above, either both state lists are of maximum size ($n/N = m/M = 1$) or there is maximal resource pool commonality and one list is of maximum size (e.g. $n/N = 1$ if $\mathcal{C} = M$). This shows that as state list size increases, favourably skewed exchange outcomes gradually disappear. The rapid rate of this decrease is illustrated in Figure 2(b).

3.1.3 Estimation of common names

Estimating the number of common names ($\mathcal{C}$) to the resource pools from exchange outcomes is of interest. If j_k, $k = 1, \cdots, K$, are K independent exchange outcomes for fixed parameters, then

$$\hat{\mathcal{C}} = \frac{2}{K\pi_0} \sum_{k=1}^{K} j_k,$$

(recall $\pi_0 = \lambda_1\lambda_2$) is an unbiased estimator of $\mathcal{C}$ and, using (3), has a standard error

$$\text{var}\, \hat{\mathcal{C}} = \frac{2}{\pi_0 K}\, \mathcal{C}\Big[1 + \frac{(\mathcal{C}-1)\pi_1}{2}\Big] - \frac{\mathcal{C}^2}{K}.$$

Moreover, an unbiased estimator of this standard error is

$$\begin{aligned} \hat{V} &:= \frac{\hat{\mathcal{C}}}{\pi_0} \frac{(2-\pi_1)}{(K + \pi_1/\pi_0 - 1)} + \hat{\mathcal{C}}^2\Big(1 - \frac{K}{K - 1 + \pi_1/\pi_0}\Big) \\ &\sim \frac{\hat{\mathcal{C}}}{K}(2/\pi_0 - 1) \quad \text{if} \quad \pi_1/\pi_0 \sim 1. \end{aligned}$$

These results area a K-sample version of the $t = 2$ estimators in section 2.3. Estimates of other parameters can be made similarly.

3.2 The 3 agent case

Consider three agents a_1, a_2, a_3 whose resource pools P_1, P_2, P_3, are described by overlap parameters C_{lm} for pairs l, m, and have $\mathcal{C}$ names in common to all three pools. The indices l, m and q are used to denote *different* agents. The $L = 3$ case, being more complex than $L = 2$, serves as an introduction to the method used for the general L case in section 4, however, more is explicitly calculable than for the general case. In particular, some joint common name distributions can be derived which in turn allow the easy computation of the variances of the exchange random variables J and J_ℓ.

3.2.1 Common name distributions

We begin by deriving the joint distribution of the number of common names between two and three state lists. Let c_{lm} be the random variable of the number of names common to the state lists of a_l and a_m, c the number of names common to all three state lists and

$d_{lm} := c_{lm} - c$ the number of names *uniquely* common to the state lists of a_l and a_m. Let r be the number of names among the $\mathcal{C}$ names common to all pools which are common to the lists of a_l and a_m, then random choice without replacement implies that

$$P(r|c_{lm}) = \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} \quad \text{and so} \quad P(r, c_{lm}) = \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} P_2(C_{lm}; c_{lm}).$$

Consequently, application of the LTP in the form $P(c|c_{lm}) = \sum_r P(c|r)P(r|c_{lm})$ gives that

$$P(c|c_{lm}) = \sum_{r=0}^{c_{lm}} \frac{\binom{r}{c}\binom{N_q-r}{n_q-c}}{\binom{N_q}{n_q}} \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}},$$

$$P(c, c_{lm}) = \sum_{r=0}^{c_{lm}} \frac{\binom{r}{c}\binom{N_q-r}{n_q-c}}{\binom{N_q}{n_q}} \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} P_2(C_{lm}; c_{lm}).$$

To check that $P(c, c_{lm})$ has the correct reductions, $\sum_c P(c, c_{lm}) = P_2(C_{lm}; c_{lm})$ is obvious, however, summing over c_{lm} requires the identity (for $\mathcal{C} \leq C_{lm}$)

$$\sum_{c_{lm}} \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} P_2(C_{lm}; c_{lm}) = P_2(\mathcal{C}; r)$$

to hold. This is just an application of the DG recursion (1) in which the parameter C_{lm} cancels and reduces the order from $t = 3$ to $t = 2$. The remaining sum over r then produces $P_3(\mathcal{C}; c)$ as per (1) so that $\sum_{c_{lm}} P(c, c_{lm}) = P_3(\mathcal{C}; c)$. From $P(c, c_{lm})$, we compute that $E[cc_{lm}] = \lambda_q(\mathcal{C}/C_{lm})E[c_{lm}^2]$, and hence,

$$E[cc_{lm}] = \lambda_l \lambda_m \lambda_q \mathcal{C}[1 + (C_{lm} - 1)\pi_1(l, m)] \tag{4}$$

follows since $E[c_{lm}^2] = \lambda_l \lambda_m C_{lm}[1 + (C_{lm} - 1)\pi_1(l, m)]$ is the second moment of the DG-distribution $P_2(C_{lm}; c_{lm})$ (see section 2.2).

The joint distribution of c_{lm} and c_{lq} can also be found by the application of the LTP and Bayes' rule. These details are relegated to an Appendix (section B.1), however, the result is:

$$P(c_{lm}, c_{lq}) = \sum_{k,r,s} \frac{\binom{k}{c_{lq}}\binom{N_q-k}{n_q-c_{lq}}}{\binom{N_q}{n_q}} \frac{\binom{s}{c_{lm}}\binom{N_m-s}{n_m-c_{lm}}}{\binom{N_m}{n_m}} \frac{\binom{C_{lq}}{k}\binom{N_l-C_{lq}}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{C_{lm}}{s}\binom{N_l-C_{lm}}{n_l-s}}{\binom{N_l}{n_l}}$$
$$\times \frac{\binom{s}{r}\binom{C_{lm}-s}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} \frac{\binom{k}{r}\binom{C_{lq}-k}{\mathcal{C}-r}}{\binom{C_{lq}}{\mathcal{C}}} \frac{\binom{N_l}{n_l}}{\binom{\mathcal{C}}{r}\binom{N_l-\mathcal{C}}{n_l-r}}. \tag{5}$$

This distribution too satisfies the appropriate reductions such as $\sum_{c_{lm}} P(c_{lm}, c_{lq}) = P_2(C_{lq}; c_{lq})$, and, when $\mathcal{C} = 0$, it factors as $P_2(C_{lm}; c_{lm})P_2(C_{lq}; c_{lq})$ to reflect the resulting independence of c_{lm} and c_{lq}. The only need for $P(c_{lm}, c_{lq})$ is in the computation of $E[c_{lm}c_{lq}]$ which is required below. Evaluation of this expectation requires the two sums $T_{lm}(r)$ and $T_{lq}(r)$ where

$$T_{lm}(r) := \sum_k k^p \frac{\binom{C_{lm}}{k}\binom{N_l-C_{lm}}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{r}\binom{B-k}{\mathcal{C}-r}}{\binom{B}{\mathcal{C}}}, \quad p = 1, B = C_{lm}.$$

The Appendix (section A.1) gives a method for evaluating this sum for p an arbitrary non-negative integer and arbitrary $B \geq \mathcal{C}$. The result for $p = 1$ and $B = C_{lm}$ (see section A.2) is:

$$T_{lm}(r) = \frac{(C_{lm} - \mathcal{C})n_l + (N_l - C_{lm})r}{N_l - \mathcal{C}} \frac{\binom{\mathcal{C}}{r}\binom{N_l - \mathcal{C}}{n_l - r}}{\binom{N_l}{n_l}}.$$

This result is the key part of the computation of $E[c_{lm}c_{lq}]$ which then turns out to be:

$$\begin{aligned} E[c_{lm}c_{lq}] = \lambda_l\lambda_m\lambda_q\Big\{&\mathcal{C}[1 + (\mathcal{C} - 1)\pi_1(l)] + \mathcal{C}\pi_1(l)(C_{lm} + C_{lq} - 2\mathcal{C}) \\ &+ (C_{lm} - \mathcal{C})(C_{lq} - \mathcal{C})\Big[\pi_1(l) + \frac{1 - \pi_1(l)}{N_l - \mathcal{C}}\Big]\Big\}. \end{aligned} \tag{6}$$

Since the joint distribution $P(c, c_{lm})$ is known, the distribution for the number of uniquely common names to two state lists; namely, $d_{lm} := c_{lm} - c$, can be written as

$$\Pr(d_{lm} = d) = \sum_{c=0}^{\mathcal{C}} P_2(C_{lm}; c + d) \sum_{r=0}^{d} \frac{\binom{r+c}{c}\binom{N-r-c}{n-c}}{\binom{N}{n}} \frac{\binom{c+d}{r+c}\binom{C_{lm}-c-d}{\mathcal{C}-r-c}}{\binom{C_{lm}}{\mathcal{C}}},$$

with the corresponding generating function

$$\begin{aligned} D_{lm}(z) &:= \sum_{c=0}^{\mathcal{C}} z^{-c} \sum_{r=c}^{C_{lm}-\mathcal{C}+c} \frac{\binom{r}{c}\binom{N-r}{n-c}}{\binom{N}{n}} P_{lm}(r; z), \quad \text{where} \\ P_{lm}(r; z) &= \sum_{d=r}^{C_{lm}-\mathcal{C}+c} P_2(C_{lm}; d) \frac{\binom{d}{r}\binom{C_{lm}-d}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} z^d. \end{aligned}$$

3.2.2 Combinatorial identities

Before leaving the common name joint distributions, it is interesting to note some combinatorial identities which follow from the LTP written in the forms

$$P(c|c_{lq}) = \sum_{c_{lm}} P(c|c_{lm})P(c_{lm}|c_{lq}) \tag{7}$$

and

$$P(c_{lm}|c_{lq}) = \sum_{c} P(c_{lm}|c)P(c|c_{lq}). \tag{8}$$

When $C_{lm} = C_{lq} = \mathcal{C}$, $P(c_{lm}, c_{lq})$ and $P(c, c_{lm})$ respectively reduce to

$$P(c_{lm}, c_{lq}) = \sum_{k=0}^{\mathcal{C}} \frac{\binom{k}{c_{lm}}\binom{N_m-k}{n_m-c_{lm}}}{\binom{N_m}{n_m}} \frac{\binom{\mathcal{C}}{k}\binom{N_l-\mathcal{C}}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{c_{lq}}\binom{N_q-k}{n_q-c_{lq}}}{\binom{N_q}{n_q}}$$

and

$$P(c, c_{lm}) = \frac{\binom{c_{lm}}{c}\binom{N_q-c_{lm}}{n_q-c}}{\binom{N_q}{n_q}} P_2(\mathcal{C}; c_{lm}).$$

For this special case, application of the forms (7) and (8) produce, respectively,

$$\sum_{k,t} \frac{\binom{t}{c}\binom{N_q-t}{n_q-c}}{\binom{N_q}{n_q}} \frac{\binom{k}{t}\binom{N_m-k}{n_m-t}}{\binom{N_m}{n_m}} \frac{\binom{\mathcal{C}}{k}\binom{N_l-\mathcal{C}}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{r}\binom{N_q-k}{n_q-r}}{\binom{N_q}{n_q}} = \frac{\binom{r}{c}\binom{N_m-r}{n_m-c}}{\binom{N_m}{n_m}} \sum_k \frac{\binom{k}{r}\binom{N_q-k}{n_q-r}}{\binom{N_q}{n_q}} \frac{\binom{\mathcal{C}}{k}\binom{N_l-\mathcal{C}}{n_l-k}}{\binom{N_l}{n_l}}$$

and

$$P_2(\mathcal{C};r)P_2(\mathcal{C};s)\sum_c \frac{\binom{r}{c}\binom{N_q-r}{n_q-c}}{\binom{N_q}{n_q}} \frac{1}{P_3(\mathcal{C};c)} \frac{\binom{s}{c}\binom{N_m-s}{n_m-c}}{\binom{N_m}{n_m}} = \sum_k \frac{\binom{k}{r}\binom{N_m-k}{n_m-r}}{\binom{N_m}{n_m}} \frac{\binom{\mathcal{C}}{k}\binom{N_l-\mathcal{C}}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{s}\binom{N_q-k}{n_q-s}}{\binom{N_q}{n_q}}$$

where $P_2(\mathcal{C};r) := P_2(\mathcal{C};n_l,n_m;N_l,N_m;r)$ and $P_2(\mathcal{C};s) := P_2(\mathcal{C};n_l,n_q;N_l,N_q;s)$. Reverting to the general case $C_{lm} \neq C_{lq} \neq \mathcal{C}$, then applying (8) and using the last identity above to replace the summation over c produces

$$\begin{aligned} P(c_{lm},c_{lq}) &= P_2(C_{lm};c_{lm})P_2(C_{lq};c_{lq}) \\ &\times \sum_{r,s,k} \frac{\binom{c_{lm}}{r}\binom{C_{lm}-c_{lm}}{\mathcal{C}-r}}{\binom{C_{lm}}{\mathcal{C}}} \frac{\binom{k}{r}\binom{N_m-k}{n_m-r}}{P_2(\mathcal{C};r)\binom{N_m}{n_m}} \frac{\binom{c_{lq}}{s}\binom{C_{lq}-c_{lq}}{\mathcal{C}-s}}{\binom{C_{lq}}{\mathcal{C}}} \frac{\binom{k}{s}\binom{N_q-k}{n_q-s}}{P_2(\mathcal{C};s)\binom{N_q}{n_q}} \frac{\binom{\mathcal{C}}{k}\binom{N_l-\mathcal{C}}{n_l-k}}{\binom{N_l}{n_l}}. \end{aligned}$$

It remains to investigate the novelty of these identities in terms of their place within existing frameworks [6, 7] .

3.2.3 The exchange problem

If j (respectively, j_ℓ) is the number of exchangeable names among the state lists of all three agents (respectively, in the state list of agent a_ℓ), write

$$\begin{aligned} j &= \sum_{l,m} j_{lm} + j_{123}, \\ j_\ell &= j_{\ell m} + j_{\ell q} + j_{123}, \end{aligned}$$

where j_{lm} is the number of exchangeable names among the d_{lm} names uniquely common to the state lists of agents a_l and a_m and j_{123} is the number of exchangeable names among those names common to all three state lists. The decomposition into exchange among uniquely common names conveniently separates common names into disjoint classes of maximum overlap among lists. The same principle is used later in the general case. From (ii) of the PL,

$$\begin{aligned} E[j] &= (1-2^{-1})\sum_{l,m} E[d_{lm}] + (1-2^{-2})E[c], \\ E[j_\ell] &= (1-2^{-1})E[d_{\ell m}] + (1-2^{-1})E[d_{\ell q}] + (1-2^{-2})E[c]. \end{aligned}$$

Since $d_{lm} = c_{lm} - c$ and $E[c_{lm}] = \lambda_l\lambda_m C_{lm}$, $E[c] = \lambda_1\lambda_2\lambda_3\mathcal{C}$, then the means $\mu = E[j]$ and $\mu_\ell = E[j_\ell]$ follow as

$$\begin{aligned} \mu &= \tfrac{1}{2}\sum_{l,m} \lambda_l\lambda_m C_{lm} - \tfrac{3}{4}\lambda_1\lambda_2\lambda_3\mathcal{C} \\ \mu_\ell &= \tfrac{1}{2}\lambda_\ell\lambda_m C_{\ell m} + \tfrac{1}{2}\lambda_\ell\lambda_q C_{\ell q} - \tfrac{1}{4}\lambda_1\lambda_2\lambda_3\mathcal{C}. \end{aligned}$$

Comparing these means with (3), the expected outcome is a sum of two-agent means less a correction for mutual overlap.

To find the corresponding variances σ^2 and σ_ℓ^2, $E[j^2]$ and $E[j_\ell^2]$ are computed using the common name joint distributions from section 3.2.1. Expanding the corresponding expressions gives

$$j^2 = \sum_{l,m} j_{lm}^2 + 2\sum_{l\neq m,q} j_{lm}j_{lq} + 2j_{123}\sum_{l,m} j_{lm} + j_{123}^2,$$
$$j_\ell^2 = j_{\ell m}^2 + j_{\ell q}^2 + 2j_{\ell m}j_{\ell q} + 2j_{123}j_{\ell m} + 2j_{123}j_{\ell q} + j_{123}^2,$$

from which, by taking expectations, it follows from (ii) of the PL that

$$E[j^2] = (1-2^{-1})^2\sum_{l,m} E[d_{lm}^2] + 2^{-1}(1-2^{-1})\sum_{l,m} E[d_{lm}] + 2(1-2^{-1})(1-2^{-1})\sum_{l\neq m,q} E[d_{lm}d_{lq}]$$
$$+ 2(1-2^{-2})(1-2^{-1})\sum_{l,m} E[cd_{lm}] + (1-2^{-2})^2E[c^2] + 2^{-2}(1-2^{-2})E[c]$$

and

$$E[j_\ell^2] = (1-2^{-1})^2E[d_{\ell m}^2] + 2^{-1}(1-2^{-1})E[d_{\ell m}] + (1-2^{-1})^2E[d_{\ell q}^2] + 2^{-1}(1-2^{-1})E[d_{\ell q}]$$
$$+ 2(1-2^{-1})(1-2^{-1})E[d_{\ell m}d_{\ell q}] + 2(1-2^{-2})(1-2^{-1})E[cd_{\ell m}]$$
$$+ 2(1-2^{-2})(1-2^{-1})E[cd_{\ell q}] + (1-2^{-2})^2E[c^2] + 2^{-2}(1-2^{-2})E[c].$$

As with the calculation of the mean, writing $d_{lm} = c_{lm} - c$ expresses $E[j^2]$ and $E[j_\ell^2]$ in terms of only the expectations involving c_{lm} and c computed in section 3.2.1. Substitution of the expectations (4) and (6) then gives

$$E[j^2] = \sum_{l,m} \tfrac{1}{2}\lambda_l\lambda_m C_{lm}[1 + \tfrac{1}{2}(C_{lm}-1)\pi_1(l,m)] + \tfrac{1}{2}\lambda_1\lambda_2\lambda_3\mathcal{C}(\mathcal{C}-1)\sum_l \pi_1(l)$$
$$- \tfrac{3}{4}\lambda_1\lambda_2\lambda_3\mathcal{C}\sum_{l,m}(C_{lm}-1)\pi_1(l,m) + \tfrac{9}{16}\lambda_1\lambda_2\lambda_3\mathcal{C}(\mathcal{C}-1)\pi_1(1,2,3) - \tfrac{3}{4}\lambda_1\lambda_2\lambda_3\mathcal{C}$$
$$+ \tfrac{1}{2}\lambda_1\lambda_2\lambda_3\Big\{\mathcal{C}\sum_{l\neq m,q}\pi_1(l)(C_{lm}+C_{lq}-2\mathcal{C}) + \sum_{l\neq m,q}(C_{lm}-\mathcal{C})(C_{lq}-\mathcal{C})[\pi_1(l) + \frac{(1-\pi_1(l))}{(N_l-\mathcal{C})}]\Big\}$$

and

$$E[j_\ell^2] = \tfrac{1}{4}\lambda_\ell\lambda_m C_{\ell m}[2 + (C_{\ell m}-1)\pi_1(\ell,m)] + \tfrac{1}{4}\lambda_\ell\lambda_q C_{\ell q}[2 + (C_{\ell q}-1)\pi_1(\ell,q)]$$
$$- \tfrac{1}{4}\lambda_1\lambda_2\lambda_3\mathcal{C}[(C_{\ell m}-1)\pi_1(\ell,m) + (C_{\ell q}-1)\pi_1(\ell,q)]$$
$$+ \tfrac{1}{16}\lambda_1\lambda_2\lambda_3\mathcal{C}(\mathcal{C}-1)\pi_1(1,2,3) - \tfrac{1}{4}\lambda_1\lambda_2\lambda_3\mathcal{C} + \tfrac{1}{2}\lambda_1\lambda_2\lambda_3\mathcal{C}(\mathcal{C}-1)\pi_1(\ell)$$
$$+ \tfrac{1}{2}\lambda_1\lambda_2\lambda_3\Big\{\mathcal{C}\pi_1(\ell)(C_{\ell m}+C_{\ell q}-2\mathcal{C}) + (C_{\ell m}-\mathcal{C})(C_{\ell q}-\mathcal{C})[\pi_1(\ell) + \frac{(1-\pi_1(\ell))}{(N_\ell-\mathcal{C})}]\Big\}.$$

Finally, the variances σ^2 and σ_ℓ^2 follow by subtracting the squares of the respective means μ and μ_ℓ.

3.2.4 Two special cases

If $\mathcal{C} = 0$, the random variables j_{lm} are independent and the means and variances simplify accordingly. Writing $\mu_{lm} = \frac{1}{2}\lambda_l\lambda_m C_{lm}$ and σ_{lm}^2 for the variance of j_{lm}, then

$$\sigma^2\Big|_{\mathcal{C}=0} = \sum_{l,m} \sigma_{lm}^2 = \sum_{l,m} \mu_{lm}[1 - \mu_{lm} + 2^{-1}(C_{lm} - 1)\pi_1(l,m)]$$

and

$$\sigma_\ell^2\Big|_{\mathcal{C}=0} = \sigma_{\ell m}^2 + \sigma_{\ell q}^2 = \mu_{\ell m}[1 - \mu_{\ell m} + 2^{-1}(C_{\ell m} - 1)\pi_1(\ell,m)] + \mu_{\ell q}[1 - \mu_{\ell q} + 2^{-1}(C_{\ell q} - 1)\pi_1(\ell,q)]$$

which expresses the variances as the sum of two agent variances (cf. (3)).

Another interesting case is when $C_{lm} = \mathcal{C}$, for all pairs l and m, so that all agents share a common 'core' of $\mathcal{C}$ resource names among their pools. In this case, the means and variances simplify somewhat with the means reducing to symmetric polynomials in the fractions λ_l (see also section 4.5). A graphical comparison with the 2-agent case can be made through further setting $n_1 = n_2 = n_3 = n$ and $N_1 = N_2 = N_3 = N$ so that all λ_l are of equal value. This describes three like agents. Then, using λ for the value of each λ_l, the means above become the cubic polynomials

$$\mu = \tfrac{3}{2}\mathcal{C}\lambda^2(1 - \lambda/2) \quad \text{and} \quad \mu_\ell = \mathcal{C}\lambda^2(1 - \lambda/4).$$

By further fixing the values for $\mathcal{C}$ and N, these means and the variances σ^2 and σ_ℓ^2, as well as the corresponding ones for the $L = 2$ case (see (3)), all become polynomials in λ. Both σ^2 and σ_ℓ^2 above are of degree six while the $L = 2$ variance is of degree four.

Figure 3(a) plots $\mu_*/\mathcal{C}$ where μ_* is each of the $L = 2$ mean $(\frac{1}{2}\lambda^2\mathcal{C})$ and μ and μ_ℓ above. As expected, the single agent mean (μ_ℓ) exceeds the 2-agent mean but falls short of the mean μ of all three agents taken together. The inflection in the curve of $\mu/\mathcal{C}$ demonstrates the onset of a characteristic "S-shape" for the system mean which becomes more pronounced with an increasing number of agents. Section 4.5 gives a more complete illustration of this behaviour.

Figure 3(b) plots $\sigma_*^2/\mathcal{C}$ where σ_*^2 is each of the $L = 2$ variance (in (3)), σ_ℓ^2 and σ^2. While for small λ each variance is an increasing function of both λ and the number of agents, the 3-agent variances exhibit a maximum and then decrease for λ approaching one. This implies that for more than two agents, there are ranges of state list size in which the exchange rate experiences a greater variance and therefore larger deviations from the mean are more probable. From the agent's viewpoint, it is desirable to operate in the reduced variance regions where smaller deviations in the exchange outcome are predicted. The extent and location of the variance's maximum can be altered by changing the parameters N and $\mathcal{C}$.

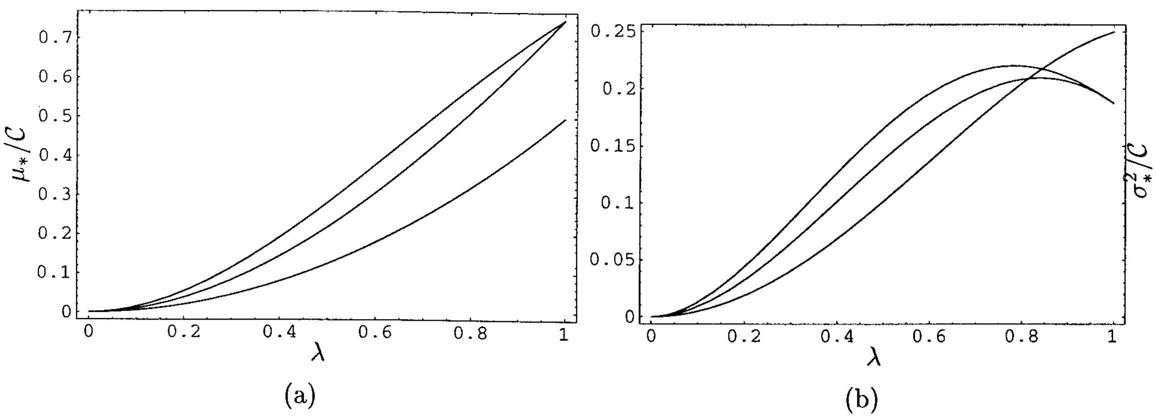


Figure 3: For the values $N = 100$ and $C = 50$, (a) plots the three means μ/C, μ_ℓ/C and $\frac{1}{2}\lambda^2$ ($L = 2$) (top to bottom) while (b) plots the variances σ^2/C and σ_ℓ^2/C as the curves exhibiting maxima ($\sigma^2 > \sigma_\ell^2$) and the variance from (3) for $L = 2$, all as functions of the state list fraction $\lambda \equiv n/N$ (see the text).

4 The general case

This section derives the mean of the number of exchangeable names for an arbitrary number L of agents. The results also include the mean for an arbitrary sub-collection (syndicate) of agents from the total of L agents. In particular, the case of all L agents considered together (the system result) and the case of a single agent (the individual result) are special cases of interest. A system administrator or resource controller would be interested in the expected exchange rate for all L agents together, while each of the independent agents is more concerned with their own personal exchange rate as a way of quickly gaining or disposing of resources. In applications, crossing network regulators and user agents are interested in achieving maximal exchange rates. The results derived here set minimum performance benchmarks for these applications.

In the next three sub-sections, all agents (section 4.1), a single agent (section 4.2) and an arbitrary subset of s of the L agents (section 4.3) are considered. Although the results in section 4.3 can be specialised to produce the single and all agents cases, the derivation in section 4.1 is important in establishing both the general method and a repeatedly used expression for the number of uniquely common names to an arbitrary collection of state lists (see (12)).

4.1 All agents

Consider now the problem of finding $\mu = E[J]$ for the whole collection of L agents. By considering all agents, evaluating μ answers the question of a system performance benchmark for the crossing network. The notation here follows that of section 1.1.

We will show that the expected number of exchangeable names μ among L unbiased

agents is

$$\mu = E[J] = \sum_{t=2}^{L}(-1)^t(1-2^{1-t}) \sum_{i_1\cdots i_t} C_{i_1\cdots i_t}\lambda_{i_1}\cdots\lambda_{i_t}. \tag{9}$$

The basic idea of the derivation is to partition the common names among state lists into disjoint sets of *uniquely* common names to k-cohorts of lists and then apply the PL. This avoids double counting the common names which lie in multiple state lists. The total number of exchangeable names is then simply the sum of the exchangeable uniquely common names (cf. section 3.2.3).

Let $x_k := \sum_{i_1\cdots i_k} d_{i_1\cdots i_k}$ denote the number of names uniquely common to all combinations of k state lists. We first show that

$$x_k = \sum_{t=k}^{L}(-1)^{t-k}\binom{t}{k}\sum_{i_1\cdots i_t} c_{i_1\cdots i_t}, \tag{10}$$

and so, by (2),

$$E[x_k] = \sum_{t=k}^{L}(-1)^{t-k}\binom{t}{k}\sum_{i_1\cdots i_t} C_{i_1\cdots i_t}\lambda_{i_1}\cdots\lambda_{i_t}. \tag{11}$$

Since the number of names uniquely common to the state lists of the agents $a_{i_1},\ldots,a_{i_k}$ is the number of common names minus the number of uniquely common names to these and all other state lists, then

$$d_{i_1\cdots i_k} = c_{i_1\cdots i_k} - \sum_{t=1}^{L-k}\sum_{\substack{j_1\cdots j_t \not\ni \\ i_1,\cdots,i_k}} d_{i_1\cdots i_k j_1\cdots j_t}.$$

Summing over all combinations $i_1\cdots i_k$ from the L agents then gives the recursion

$$x_k = \sum_{i_1\cdots i_k} c_{i_1\cdots i_k} - \sum_{t=1}^{L-k}\tbinom{t+k}{k}x_{t+k},$$

where the identity (a) from the CL has been used. By (i) of the Lemma (section B.2), the solution of the recursion is just (10).

Knowing x_k permits the derivation of an important formula for $d_{i_1\cdots i_k}$ which is used repeatedly below. For unknown coefficients α_t, write

$$d_{i_1\cdots i_k} = \sum_{t=0}^{L-k}\alpha_t\sum_{\substack{j_1\cdots j_t \not\ni \\ i_1,\cdots,i_k}} c_{i_1\cdots i_k j_1\cdots j_t}$$

Summing over $i_1\cdots i_k$ shows that $x_k = \sum_{t=k}^{L}\alpha_{t-k}\binom{t}{k}\sum_{i_1\cdots i_t} c_{i_1\cdots i_t}$. Comparing this with x_k in (10) then gives $\alpha_t = (-1)^t$, and consequently,

$$d_{i_1\cdots i_k} = \sum_{t=0}^{L-k}(-1)^t\sum_{\substack{j_1\cdots j_t \not\ni \\ i_1,\cdots,i_k}} c_{i_1\cdots i_k j_1\cdots j_t}. \tag{12}$$

This formula, which expresses the number of uniquely common names in terms of the number of common names, is the main tool in many of the derivations which follow.

Let $j_{i_1 \cdots i_k}$ be the number of exchangeable names in the $d_{i_1 \cdots i_k}$ names uniquely common to the state lists of $a_{i_1}, \ldots, a_{i_k}$. Then the total number of exchangeable names is

$$j = \sum_{k=2}^{L} \sum_{i_1 \cdots i_k} j_{i_1 \cdots i_k}.$$

Since, by (ii) of the PL, $E[j_{i_1 \cdots i_k}] = (1 - 2^{1-k}) E[d_{i_1 \cdots i_k}]$, then $\sum_{i_1 \cdots i_k} E[j_{i_1 \cdots i_k}] = (1 - 2^{1-k}) E[x_k]$ and so, by (11),

$$\mu = E[j] = \sum_{k=2}^{L} (1 - 2^{1-k}) \sum_{t=k}^{L} (-1)^{t-k} \binom{t}{k} \sum_{i_1 \cdots i_t} C_{i_1 \cdots i_t} \lambda_{i_1} \cdots \lambda_{i_t}.$$

Reversing the order of the t and k summations finally gives the result in (9).

The expression for μ can be partitioned to illustrate the *marginal increase* by adding one extra agent. If $\mu(L)$ temporarily denotes the mean for L agents, then (9) implies that the marginal increase in the expected exchange rate of all agents in going from $L-1$ agents to L agents through the addition of agent a_L is:

$$\mu(L) - \mu(L-1) = \lambda_L \sum_{t=2}^{L} (-1)^t (1 - 2^{1-t}) \sum_{i_1 \cdots i_{t-1}} C_{i_1 \cdots i_{t-1} L} \lambda_{i_1} \cdots \lambda_{i_{t-1}}. \tag{13}$$

The expressions (9) and (13) are studied further in section 4.5 below.

It is also possible to derive μ via a similar argument but based upon counting the number of exchangeable names among those uniquely common to resource pools rather than state lists. This alternative derivation is given in section B.3.

4.2 A single agent

This section solves the problem of evaluating $\mu_\ell = E[J_\ell]$ for the specific agent a_ℓ. We show that the expected number of exchangeable names μ_ℓ in the state list of agent a_ℓ is

$$\mu_\ell = E[J_\ell] = \lambda_\ell \sum_{t=1}^{L-1} (-1)^{t+1} 2^{-t} \sum_{i_1 \cdots i_t \not\ni \ell} C_{\ell i_1 \cdots i_t} \lambda_{i_1} \cdots \lambda_{i_t}. \tag{14}$$

The idea of the derivation of μ_ℓ is the same as for μ except that the single index ℓ is frozen during the process. If $x_{\ell,k} := \sum_{i_1 \cdots i_{k-1}} d_{\ell i_1 \cdots i_{k-1}}$ denotes the number of names uniquely common to k state lists which include the list of a_ℓ, we first show that

$$x_{\ell,k} = \sum_{t=k-1}^{L-1} (-1)^{t-k+1} \binom{t}{k-1} \sum_{i_1 \cdots i_t} c_{\ell i_1 \cdots i_t}, \tag{15}$$

and so, by (2),

$$E[x_{\ell,k}] = \lambda_\ell \sum_{t=k-1}^{L-1} (-1)^{t-k+1} \binom{t}{k-1} \sum_{i_1 \cdots i_t} C_{\ell i_1 \cdots i_t} \lambda_{i_1} \cdots \lambda_{i_t}. \tag{16}$$

By fixing one index as ℓ in (12),

$$d_{\ell i_1 \cdots i_{k-1}} = \sum_{t=0}^{L-k} (-1)^t \sum_{\substack{j_1 \cdots j_t \not\ni \\ \ell, i_1, \cdots, i_{k-1}}} c_{\ell i_1 \cdots i_{k-1} j_1 \cdots j_t}$$

so that summing over $i_1 \cdots i_{k-1}$ with the use of (a) of the CL produces (15). Note that since $\sum_{k=2}^{L} x_{\ell,k}$ is the number of names in the state list of agent a_ℓ in common with some other state list and therefore cannot exceed the state list's size, then

$$\sum_{k=2}^{L} x_{\ell,k} = \sum_{t=1}^{L-1} (-1)^{t+1} \sum_{i_1 \cdots i_t} c_{\ell i_1 \cdots i_t} \le n_\ell$$

always holds.

Let $j_{\ell i_1 \cdots i_{k-1}}$ be the number of exchangeable names in the $d_{\ell i_1 \cdots i_{k-1}}$ names uniquely common to the state lists of a_ℓ, $a_{i_1}, \ldots, a_{i_{k-1}}$. Then the total number of exchangeable names in the state list of a_ℓ is

$$j_\ell = \sum_{k=2}^{L} \sum_{i_1 \cdots i_{k-1} \not\ni \ell} j_{\ell i_1 \cdots i_{k-1}}.$$

Again by (ii) of the PL, $E[j_{\ell i_1 \cdots i_{k-1}}] = (1 - 2^{1-k}) E[d_{\ell i_1 \cdots i_{k-1}}]$ so that summing over the combinations gives $\sum_{i_1 \cdots i_{k-1}} E[j_{\ell i_1 \cdots i_{k-1}}] = (1 - 2^{1-k}) E[x_{\ell,k}]$. Hence, by (16),

$$\mu_\ell = E[j_\ell] = \sum_{k=2}^{L} (1 - 2^{1-k}) \sum_{t=k}^{L} (-1)^{t-k} \binom{t-1}{k-1} \sum_{i_1 \cdots i_{t-1} \not\ni \ell} C_{\ell i_1 \cdots i_{t-1}} \lambda_\ell \lambda_{i_1} \cdots \lambda_{i_{t-1}}.$$

Reversing the order of the t and k summations gives the result in (14).

As with all agents, the *marginal increase* in the mean μ_ℓ can be similarly deduced from (14). If $\mu_\ell(L)$ denotes the mean (14) for L agents present, then the marginal increase in the expected exchange rate for agent a_ℓ in going from $L-1$ agents to L agents through the addition of agent a_L is:

$$\mu_\ell(L) - \mu_\ell(L-1) = \lambda_\ell \lambda_L \sum_{t=1}^{L-1} (-1)^{t+1} 2^{-t} \sum_{i_1 \cdots i_{t-1} \not\ni \ell, L} C_{\ell L i_1 \cdots i_{t-1}} \lambda_{i_1} \cdots \lambda_{i_{t-1}}. \tag{17}$$

Again, both expressions (14) and (17) are studied in further detail in section 4.5.

4.3 A syndicate of agents

Consider now the (interpolating) problem of finding the expected number of exchangeable names within the state lists of a specified syndicate $a_{I_s} \equiv a_{\ell_1}, \ldots, a_{\ell_s}$, $2 \leq s \leq L-1$, of the L agents. By re-numbering the indices, it can be assumed that $\ell_1, \ldots, \ell_s$ are the first s agents. Let $I_s := \{\ell_1, \ldots, \ell_s\}$ be these indices and $I \backslash I_s$ be the remaining indices. As in section 1.1, let J_{I_s} be the random variable of the number of exchangeable names within the state lists of the agents a_{I_s} and let $\mu_{I_s} = E[J_{I_s}]$ be its mean.

The number j_{I_s} of exchangeable names can be written in two parts as:

$$\begin{aligned} j_{I_s} &= \sum_{p=2}^{s} \sum_{\substack{r_1 \cdots r_p \\ \in I_s}} j_{r_1 \cdots r_p} + \sum_{p=1}^{s} \sum_{k=1}^{L-s} \sum_{\substack{r_1 \cdots r_p \\ \in I_s}} \sum_{\substack{i_1 \cdots i_k \\ \in I \backslash I_s}} j_{r_1 \cdots r_p i_1 \cdots i_k}, \\ &=: A + B \end{aligned}$$

which reflects a decomposition into the exchangeable names in p-cohorts taken only from I_s (i.e. A) and those in $(k+p)$-cohorts containing at least one agent from I_s and one agent from $I \backslash I_s$ (i.e. B). To find μ_{I_s}, both $E[A]$ and $E[B]$ must be computed. Consider A first. Since by (ii) of the PL, $E[A] = \sum_{p=2}^{s}(1-2^{1-p}) \sum_{r_1 \cdots r_p} E[d_{r_1 \cdots r_p}]$, the rewriting of the sum

$$d_{r_1 \cdots r_p} = \sum_{t'=0}^{L-p} (-1)^{t'} \sum_{\substack{j_1 \cdots j_{t'} \\ \not\ni r_1, \cdots, r_p}} c_{r_1 \cdots r_p j_1 \cdots j_{t'}}$$

as

$$= \sum_{q=0}^{s-p} \sum_{t=0}^{L-s} (-1)^{q+t} \sum_{\substack{l_1 \cdots l_q \\ \in I_s \backslash \{r_1, \cdots, r_p\}}} \sum_{\substack{j_1 \cdots j_t \\ \in I \backslash I_s}} c_{r_1 \cdots r_p l_1 \cdots l_q j_1 \cdots j_t}$$

splits the summation over the $j_{t'}$ into two *disjoint* pieces, one within I_s and one within $I \backslash I_s$. Summing over the r_p's and taking the expectation now gives

$$\begin{aligned} E[A] &= \sum_{p=2}^{s} (1-2^{1-p}) \sum_{q=0}^{s-p} \sum_{t=0}^{L-s} (-1)^{q+t} \sum_{r_1 \cdots r_p} \sum_{l_1 \cdots l_q} \sum_{j_1 \cdots j_t} E[c_{r_1 \cdots r_p l_1 \cdots l_q j_1 \ldots j_t}] \\ &= \sum_{p=2}^{s} (1-2^{1-p}) \sum_{q=p}^{s} \sum_{t=0}^{L-s} (-1)^{t+q-p} \binom{q}{p} \sum_{r_1 \cdots r_q} \sum_{j_1 \cdots j_t} E[c_{r_1 \cdots r_q j_1 \cdots j_t}] \end{aligned}$$

where the CL and a shift of p in the q-summation have been used. Next, switching the order of the p and q summations and using $\sum_{p=1}^{q}(1-2^{1-p})(-1)^p \binom{q}{p} = 1 - 2^{1-q}$ gives

$$E[A] = \sum_{t=0}^{L-s} \sum_{p=2}^{s} (-1)^{t+q} (1-2^{1-q}) \sum_{\substack{r_1 \cdots r_q \\ \in I_s}} \sum_{\substack{i_1 \cdots i_t \\ \in I \backslash I_s}} C_{r_1 \cdots r_q i_1 \cdots i_t} \lambda_{r_1} \cdots \lambda_{r_q} \lambda_{i_1} \cdots \lambda_{i_t}.$$

Now consider B. As for A, again write

$$d_{r_1\cdots r_p i_1\cdots i_k} = \sum_{t'=0}^{L-p-k} (-1)^{t'} \sum_{\substack{j_1\cdots j_{t'} \\ \not\ni r_1,\cdots,i_k}} c_{r_1\cdots r_p i_1\cdots i_k j_1\cdots j_{t'}}$$
$$= \sum_{q=0}^{s-p} \sum_{t=0}^{L-s-k} (-1)^{q+t} \sum_{\substack{l_1\cdots l_q \\ \not\ni r_1,\dots,r_p}} \sum_{\substack{j_1\cdots j_t \\ \not\ni i_1,\dots,i_k}} c_{r_1\cdots r_p i_1\cdots i_k l_1\cdots l_q j_1\cdots j_t}$$

to split the j-summation over I_s and $I\backslash I_s$. Since by (a) of the CL,

$$\sum_{r_1\cdots r_p} \sum_{\substack{l_1\cdots l_q \\ \not\ni r_1,\dots,r_p}} \sum_{i_1\cdots i_k} \sum_{\substack{j_1\cdots j_t \\ \not\ni i_1,\dots,i_k}} c_{r_1\cdots r_p i_1\cdots i_k l_1\cdots l_q j_1\cdots j_t} =$$
$$\binom{q+p}{p}\binom{t+k}{k} \sum_{r_1\cdots r_{p+q}} \sum_{i_1\cdots i_{k+t}} c_{r_1\cdots r_{q+p} i_1\cdots i_{k+t}}$$

substitution of the expectation of $d_{r_1\cdots r_p i_1\cdots i_k}$ into $E[B]$ gives

$$\sum_{p=1}^{s}\sum_{k=1}^{L-s}(1-2^{1-p-k})\sum_{q=p}^{s}\sum_{t=k}^{L-s}(-1)^{t-k+q-p}\binom{q}{p}\binom{t}{k}\sum_{r_1\cdots r_q}\sum_{i_1\cdots i_t} E[c_{r_1\cdots r_q i_1\cdots i_t}],$$

where the q and t summations have been shifted by p and k respectively. Interchanging the p and q summations and k and t summations then produces

$$\sum_{q=1}^{s}\sum_{t=1}^{L-s}(-1)^{t+q}\left[\sum_{p=1}^{q}(-1)^p\binom{q}{p}\sum_{k=1}^{t}(-1)^k\binom{t}{k}(1-2^{1-p-k})\right]\sum_{r_1\cdots r_q}\sum_{i_1\cdots i_t} E[c_{r_1\cdots r_q i_1\cdots i_t}].$$

Evaluating the sum in the inner bracket as $1-2(1-2^{-t})(1-2^{-q})$ then gives

$$E[B] = \sum_{q=1}^{s}\sum_{t=1}^{L-s}(-1)^{t+q}[1-2(1-2^{-t})(1-2^{-q})]\sum_{r_1\cdots r_q}\sum_{i_1\cdots i_t} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}.$$

Finally, combining the results for $E[A]$ and $E[B]$ produces the result

$$\mu_{I_s} = \sum_{q=2}^{s}\sum_{t=0}^{L-s}(-1)^{t+q}(1-2^{1-q})\sum_{\substack{r_1\cdots r_q \\ \in I_s}}\sum_{\substack{i_1\cdots i_t \\ \in I\backslash I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}$$
$$+\sum_{q=1}^{s}\sum_{t=1}^{L-s}(-1)^{t+q}[1-2(1-2^{-t})(1-2^{-q})]\sum_{\substack{r_1\cdots r_q \\ \in I_s}}\sum_{\substack{i_1\cdots i_t \\ \in I\backslash I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}$$

or, equivalently,

$$\mu_{I_s} = \sum_{q=2}^{s}(-1)^q(1-2^{1-q}) \sum_{\substack{r_1\cdots r_q \\ \in I_s}} C_{r_1\cdots r_q}\lambda_{r_1}\cdots\lambda_{r_q}$$

$$+\sum_{q=1}^{s}\sum_{t=1}^{L-s}(-1)^{t+q}2^{1-t}(1-2^{-q}) \sum_{\substack{r_1\cdots r_q \\ \in I_s}} \sum_{\substack{i_1\cdots i_t \\ \in I\setminus I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}. \tag{18}$$

Written in this way, μ_{I_s} is the sum of the exchange rate for the agents a_{I_s} considered in isolation (cf. 9) plus a correction term accounting for the interactions with the remaining agents not indexed by I_s. If the agents a_{I_s} form a *cooperative syndicate* in the sense that they share minimal interest in common resources, i.e. $C_{r_1\cdots r_p} = 0$ for all combinations $r_1\cdots r_p \in I_s$, then μ_{I_s} in (18) consists only of the second interaction term.

As a final check on the above results, we verify the means computed in the previous two sections. First, for the single agent a_ℓ ($I_s = \{\ell\}$), set $s = 1$. In this case $E[A]$ must be dropped from μ_ℓ and by setting $q = 1$, its only possible value in $E[B]$, gives

$$\mu_\ell = \sum_{t=1}^{L-1}(-1)^{t+1}2^{-t} \sum_{\substack{i_1\cdots i_t \\ \not\ni \ell}} C_{\ell i_1\cdots i_t}\lambda_\ell\lambda_{i_1}\cdots\lambda_{i_t},$$

which is exactly the previous result for a single agent in (14). Next, for all L agents, set $s = L$ (i.e. $I_s = \{1,\ldots,L\}$). In this case $E[B]$ must be dropped from μ (no interaction) and setting $t = 0$ in $E[A]$ produces

$$\mu = \sum_{q=2}^{L}(-1)^q(1-2^{1-q}) \sum_{r_1\cdots r_q} C_{r_1\cdots r_q}\lambda_{r_1}\cdots\lambda_{r_q},$$

which is the same as derived previously for the system mean (9). Hence, the general result for μ_{I_s} encapsulates the two previous cases ($s = 1$ and $s = L$).

4.4 The number of common names

A similar analysis to the previous section applies for the number of common names to the state lists of the agents indexed by I_s. Let x_{I_s} be the random variable of the number of names in the state lists of agents a_{I_s} common to two or more state lists and let $\chi_{I_s} := E[x_{I_s}]$ be its mean, then

$$\begin{aligned} x_{I_s} &= \sum_{p=2}^{s}\sum_{\substack{r_1\cdots r_p \\ \in I_s}} d_{r_1\cdots r_p} + \sum_{p=1}^{s}\sum_{k=1}^{L-s}\sum_{\substack{r_1\cdots r_p \\ \in I_s}}\sum_{\substack{i_1\cdots i_k \\ \in I\setminus I_s}} d_{r_1\cdots r_p i_1\cdots i_k} \\ &=: a + b. \end{aligned}$$

Since x_{I_s} is a count of the number of names appearing in the state lists of agents a_{I_s} which are common to two or more state lists, it tallies the number of names in which exchange

is possible. To find the expected value χ_{I_s}, the computation of $E[a]$ and $E[b]$ is required. A very similar derivation to that of μ_{I_s} above, but one that does not use the PL, shows that

$$E[a] = \sum_{q=2}^{s}\sum_{t=0}^{L-s}(-1)^{t+q}(q-1)\sum_{\substack{r_1\cdots r_q\\ \in I_s}}\sum_{\substack{i_1\cdots i_t\\ \in I\setminus I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t},$$

$$E[b] = \sum_{q=1}^{s}\sum_{t=1}^{L-s}(-1)^{t+q}\sum_{\substack{r_1\cdots r_q\\ \in I_s}}\sum_{\substack{i_1\cdots i_t\\ \in I\setminus I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}.$$

Consequently, the expected number of names in the state lists of agents a_{I_s} which are common to two or more state lists is

$$\begin{aligned}\chi_{I_s} &= \sum_{q=2}^{s}(-1)^q(q-1)\sum_{\substack{r_1\cdots r_q\\ \in I_s}} C_{r_1\cdots r_q}\lambda_{r_1}\cdots\lambda_{r_q}\\ &+\sum_{q=1}^{s}\sum_{t=1}^{L-s}q(-1)^{t+q}\sum_{\substack{r_1\cdots r_q\\ \in I_s}}\sum_{\substack{i_1\cdots i_t\\ \in I\setminus I_s}} C_{r_1\cdots r_q i_1\cdots i_t}\lambda_{r_1}\cdots\lambda_{r_q}\lambda_{i_1}\cdots\lambda_{i_t}.\end{aligned} \tag{19}$$

Again, this expression is split as the sum of the expected number of common names among the agents a_{I_s} as if in isolation plus a correction term to account for the interactions with the agents not indexed by I_s.

One use of the expression for χ_{I_s} is to obtain upper bounds on the possible realisations of the random variables J, J_ℓ and J_{I_s}. This is done by taking all of the state lists of maximum size in which case each is equal to the corresponding resource pool. For $s = L$, the term $E[b]$ must be dropped to obtain an expression for the expected number of common names among all state lists (χ). Setting $t = 0$ in $E[a]$ and every $\lambda_{i_k} \equiv 1$ to give maximal state list size produces

$$\chi_{\max} = \sum_{q=2}^{L}(-1)^q(q-1)\sum_{r_1\cdots r_q} C_{r_1\cdots r_q}.$$

$\chi_{\max}$ is the number of names common to two or more resource pools and as such is the maximum possible number of exchangeable names among all L state lists. In other words, any realisation j of J must satisfy $j \le \chi_{\max}$.

For $s = 1$, the term $E[a]$ must be dropped from the expression for χ_ℓ. Setting $q = 1$ in $E[b]$ and every $\lambda_{i_k} \equiv 1$ to give maximum state list size produces

$$\chi_{\ell,\max} = \sum_{t=1}^{L-1}(-1)^{t+1}\sum_{i_1\cdots i_t \not\ni \ell} C_{\ell i_1\cdots i_t}.$$

$\chi_{\ell,\max}$ is the number of names in the pool P_ℓ which are common to two or more resource pools and is therefore the maximum possible number of exchangeable names that can ever appear in the state list of agent a_ℓ. That is, any realisation j_ℓ of J_ℓ must satisfy $j_\ell \le \chi_{\ell,\max}$.

For the intermediate values, $2 \leq s \leq L-1$, and with $I_s = \{\ell_1, \cdots, \ell_s\}$, putting every $\lambda_{i_k} \equiv 1$ in (19) produces

$$\chi_{I_s,\max} = \sum_{q=2}^{s}(-1)^q(q-1)\sum_{\substack{r_1\cdots r_q\\ \in I_s}} C_{r_1\cdots r_q} + \sum_{q=1}^{s}\sum_{t=1}^{L-s} q(-1)^{t+q}\sum_{\substack{r_1\cdots r_q\\ \in I_s}}\sum_{\substack{i_1\cdots i_t\\ \in I\setminus I_s}} C_{r_1\ldots r_q i_1\cdots i_t}$$

as the number of names in the pools P_{I_s} which are common to two or more pools. $\chi_{I_s,\max}$ is therefore an upper bound for any realisation j_{I_s} of J_{I_s} in that $j_{I_s} \leq \chi_{I_s,\max}$ must always hold.

4.5 Reductions under EDCI

To easily visualise the behaviour of the mean with the number of agents L, the relative state list size and the number of the selected agents a_{I_s}, it is necessary to sufficiently collapse the high dimensional dependence in the parameters $C_{i_1\cdots i_t}$ and λ_{i_t}. This is done using the EDCI assumption (section 1.1) which, by giving a diminishing common resource interest with increasing numbers of agents, also allows closed form evaluations of the multi-index sums involved and a subsequent reduction to one-dimensional polynomials.

The EDCI substitution $C_{i_1\cdots i_t} = \rho^t \mathcal{C}$, $0 \leq \rho \leq 1$, reduces the expressions for μ_{I_s} in (18) and χ_{I_s} in (19) to simple (symmetric) polynomials in the variables $\rho\lambda_l$. For a given index set I_s, define the quantities

$$X = 1 - \prod_{l\in I\setminus I_s}(1-\rho\lambda_l), \qquad X_s = 1 - \prod_{l\in I_s}(1-\rho\lambda_l),$$

$$Y = 1 - \prod_{l\in I\setminus I_s}(1-\rho\lambda_l/2), \qquad Y_s = 1 - \prod_{l\in I_s}(1-\rho\lambda_l/2),$$

$$Z_s = \sum_{l\in I_s}\rho\lambda_l \prod_{m\in I_s\setminus\{l\}}(1-\rho\lambda_m).$$

First consider the mean μ_{I_s}. Substitution for $C_{i_1\cdots i_t}$ into (18) and evaluating the resulting sums shows that

$$\mu_{I_s} = \mathcal{C}\Big[(1-X)(2Y_s - X_s) + XX_s - 2(Y-X)(Y_s - X_s)\Big], \quad 1 \leq s \leq L,$$

where, although derived only for $2 \leq s \leq L-1$, the expression remains valid for $1 \leq s \leq L$. Now consider the expected number of common names χ_{I_s}. Substitution for $C_{i_1\cdots i_t}$ into (19) produces

$$\chi_{I_s} = \mathcal{C}\Big[X_s + (X-1)Z_s\Big], \quad 1 \leq s \leq L.$$

This expression is also valid for the full range of values of s. Observe that the sizes of the corresponding pool overlaps, as computed in section 4.4, are obtained by letting $\lambda_l \to 1$, for all l, and produces

$$\begin{aligned}
\chi_{\max} &= [1-(1-\rho)^{L-1} - \rho L(1-\rho)^{L-1}]\mathcal{C},\\
\chi_{\ell,\max} &= \rho[1-(1-\rho)^{L-1}]\mathcal{C},\\
\chi_{I_s,\max} &= [1-(1-\rho)^{s} - s\rho(1-\rho)^{L-1}]\mathcal{C}.
\end{aligned}$$

As expected, each of these satisfies $\lim_{\rho\to 1}\chi_{\cdot,\max} = \mathcal{C}$ for the case of a 'core' of $\mathcal{C}$ common resources and $\lim_{\rho\to 0}\chi_{\cdot,\max} = 0$ for the case of no common resources.

The efficiency $e_{I_s} := \mu_{I_s}/\chi_{I_s}$ defined earlier in section 1.1 can now be written as

$$e_{I_s} = \frac{(1-X)(2Y_s - X_s) + XX_s - 2(Y-X)(Y_s - X_s)}{X_s + (X-1)Z_s}.$$

To produce one-dimensional plots, consider like agents for whom $\lambda_l = \lambda$, for all l, so that

$$\begin{aligned}
X &= 1-(1-\rho\lambda)^{L-s}, & X_s &= 1-(1-\rho\lambda)^{s},\\
Y &= 1-(1-\rho\lambda/2)^{L-s}, & Y_s &= 1-(1-\rho\lambda/2)^{s},\\
Z_s &= s\rho\lambda(1-\rho\lambda)^{s-1}, & &
\end{aligned}$$

are all polynomials in $\rho\lambda$.

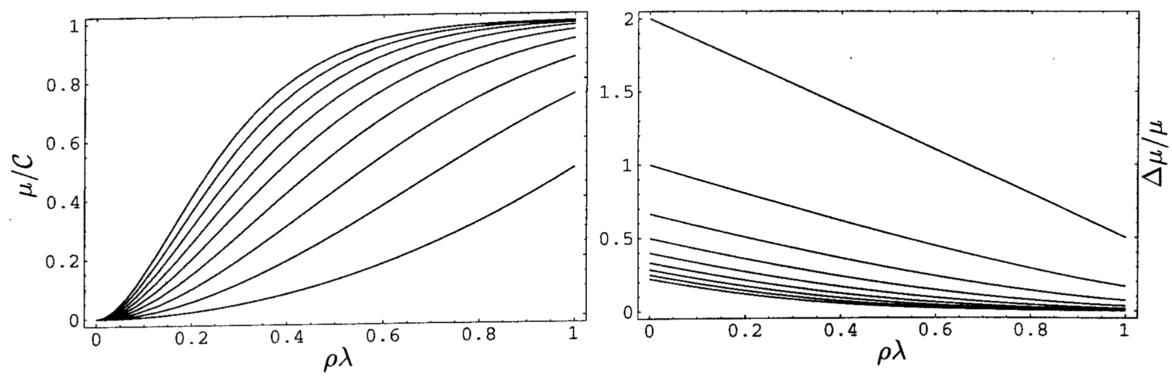
For $s = L$, $I_L = \{1,\ldots,L\}$ designates all agents and the values $X = Y = 0$ should be used in the above expressions. Hence, $\mu = (2Y_L - X_L)\mathcal{C}$ and $\chi = (X_L - Z_L)\mathcal{C}$. If, as in section 4.1, $\mu(L-1)$ denotes the mean number of exchangeable names for $L-1$ agents in isolation, then the relative marginal increase by adding agent a_L is

$$\frac{\Delta_L\mu(L-1)}{\mu(L-1)} = \rho\lambda_L\,\frac{X_{L-1} - Y_{L-1}}{2Y_{L-1} - X_{L-1}}.$$

For the further reduction $\lambda_l = \lambda$, for all l, Figure 4(a) plots the mean $\mu/\mathcal{C}$ while Figure 4(b) plots the relative margin $\Delta_L\mu(L-1)/\mu(L-1)$, both as functions of $\rho\lambda$ for various values of L. Whereas the mean increases, the relative margin decreases as a function of both $\rho\lambda$ and L. Figure 5(a) plots the efficiency $e = (2Y_L - X_L)/(X_L - Z_L)$ for various values of L as a function of $\rho\lambda$. Figure 5(b) plots efficiency contours in $\rho\lambda$-L space. These are generated by, for each efficiency level ϵ, solving $e(\rho\lambda) = \epsilon$ for $\rho\lambda$ for each fixed L. Note that certain high levels of efficiency are not attained without sufficiently many agents as is indicated by the failure of the corresponding curves in Figure 5(a) to ever reach those efficiency values.

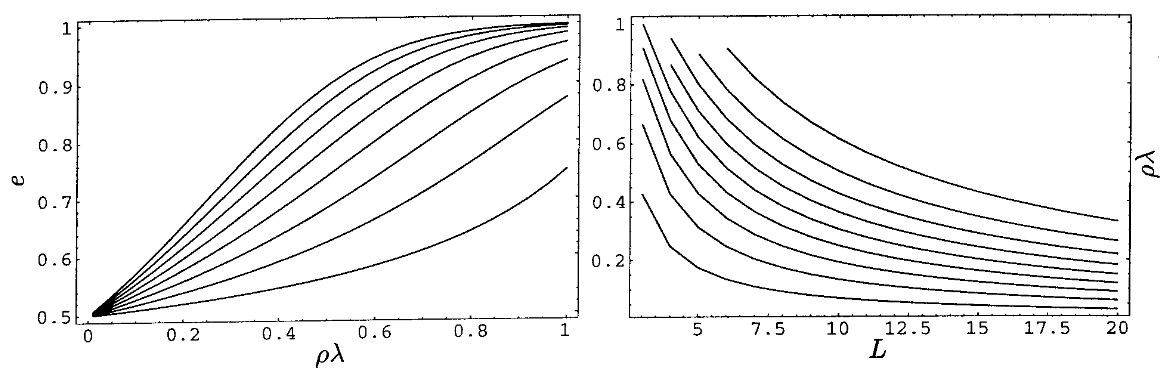
For $s = 1$, $I_1 = \{\ell\}$ designates the single agent a_ℓ and $Z_1 = X_1 = 2Y_1 = \rho\lambda_\ell$. In this case, the mean is $\mu_\ell = \rho\lambda_\ell Y\mathcal{C}$ and the expected number of common names is $\chi_\ell = \rho\lambda_\ell X\mathcal{C}$. If, as in section 4.2, $\mu_\ell(L-1)$ denotes the mean for agent a_ℓ among $L-1$ agents in isolation, then the relative marginal increase through adding one more agent a_L is

$$\frac{\Delta_L\mu_\ell(L-1)}{\mu_\ell(L-1)} = \tfrac{1}{2}\rho\lambda_L\,\frac{1-Y(L-1)}{Y(L-1)},$$



(a) (b)

Figure 4: (a) A plot of μ/C for $L = 2, \ldots, 10$ (bottom to top); (b) plot of the relative margin $\Delta_L\mu(L-1)/\mu(L-1)$ versus $\rho\lambda$ for $L = 3, \ldots, 11$ (top to bottom).



(a) (b)

Figure 5: (a) A plot of the efficiency $e(\rho\lambda)$ for $L = 3, \ldots, 10$ (bottom to top); (b) efficiency contours in L-$\rho\lambda$ space for $\epsilon = 0.95, 0.90, \ldots, 0.55$ (top to bottom).

where $Y(L-1)$ denotes Y defined only over the agents $a_1, \ldots, a_{L-1}$. Figure 6(a) plots the ratio $\mu_\ell/(\rho\lambda_\ell C)$ and Figure 6(b) plots the relative margin $\Delta_L\mu_\ell(L-1)/\mu_\ell(L-1)$, for various values of L, as functions of $\rho\lambda$ under the further reduction $\lambda_m = \lambda$, for all $m \neq \ell$. Again the mean increases and the relative margin decreases as functions of both $\rho\lambda$ and L. The efficiency $e_\ell = Y/X$ is very similar in shape to e as a function of both $\rho\lambda$ and L. However, $e_\ell > e$, and if $\lambda_m = \lambda_\ell = \lambda$ for all m, then the peak difference remains bounded according to $0.02 \leq \max_L \max_{\rho\lambda}(e_\ell - e) \leq 0.06$ as illustrated in Figure 7.

Next consider the convergence of μ_{I_s} to μ and e_{I_s} to e as s increases from 1 to L. Again by putting $\lambda_l = \lambda$, for all l, Figure 8 shows plots of the mean μ_{I_s} in (a) and the efficiency e_{I_s} in (b) for $L = 10$ and $s = 1, \ldots, 10$. Observe that while μ_{I_s} increases from the individual mean μ_ℓ to the system mean μ with increasing s, e_{I_s} decreases with s from the individual efficiency e_ℓ to its system value e for all L agents. Hence, when considered as a group, any syndicate of s agents enjoys a higher efficiency but a lower mean than the whole collection with monotone convergence to the L-agent system values as s increases.

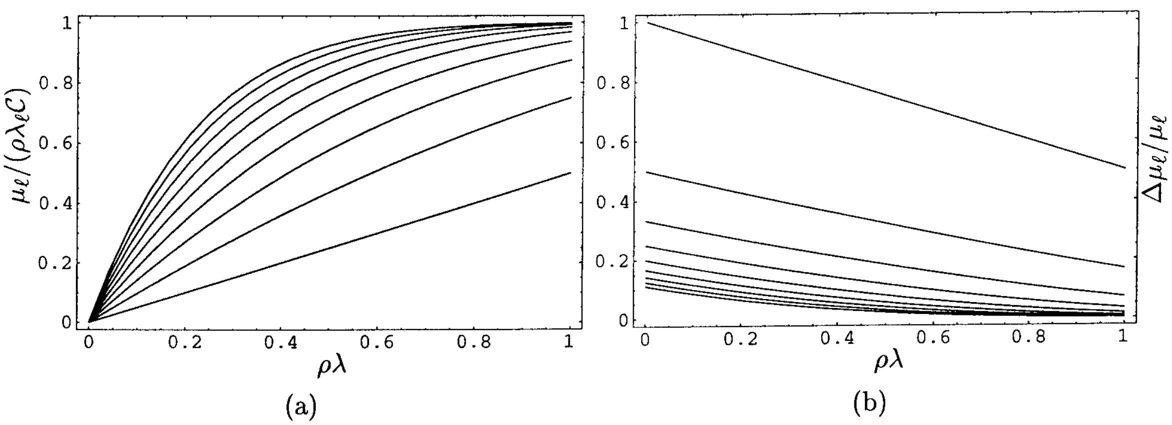


Figure 6: (a) A plot of $\mu_\ell/(\rho\lambda_\ell C)$ (bottom to top); (b) the single agent relative margin $\Delta_L\mu_\ell(L-1)/\mu_\ell(L-1)$ as functions of $\rho\lambda$ for $L = 3, \ldots, 10$ (top to bottom).

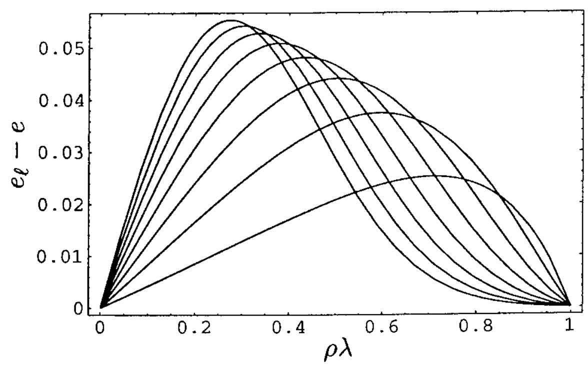


Figure 7: A plot of the efficiency difference $e_\ell - e$ as a function of $\rho\lambda$ for the values $L = 3, \ldots, 10$.

Figures 4, 5, 7 and 8 are indicative performance curves for networks of like agents. Of course, if the many parameters are known, the exact formulae from section 4 can be used, however, the simple one-dimensional plots here provide intuition as to how performance should change with both the number of agents present and the fractions of their pools represented within state lists.

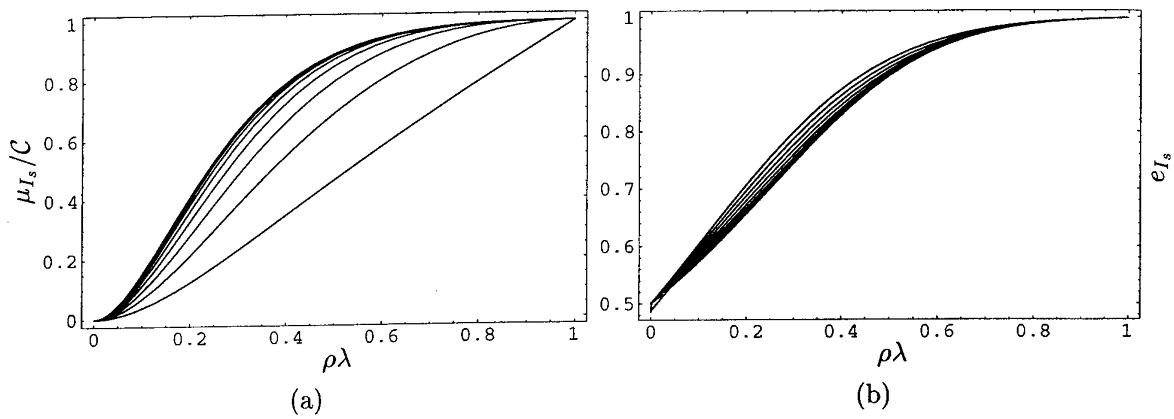


Figure 8: Plots of the means μ_{I_s} in (a) and the efficiencies e_{I_s} in (b) as functions of $\rho\lambda$ for the values $s = 1, \ldots, 10$ with $L = 10$.

5 Conclusion

This report has developed a combinatorial probability model for the expected exchange rate of resources when $L \geq 2$ unbiased agents meet in a crossing network. The unbiased assumption has allowed the derivation of explicit formulae for, in simple cases ($L = 2, 3$), the exchange distribution and certain higher moments of the distribution. In particular, for $L = 2$ this distribution is favourably (positively) skewed, but only significantly so for small values of the mean and variance, a property which rapidly diminishes as state list size increases.

For general L, the mean of this distribution (i.e. the exchange rate), has been found for all of the agents, a single agent and an arbitrary syndicate of agents. Through a suitable reduction of the number of defining parameters (EDCI), reasonable intuition can be gained for the behaviour of crossing networks through the development of one-dimensional expected outcome plots as were detailed in section 4.5. For example, while the means μ and μ_ℓ are increasing functions of L and state list size, the relative margins $\Delta_L\mu/\mu$ and $\Delta_L\mu_\ell/\mu_\ell$ are both decreasing functions. The individual efficiency e_ℓ exceeds the system efficiency e, however, both are increasing functions of relative state list size. Such intuition (see Figures 4 - 8) is useful for administrators, regulators and users of crossing networks for both the prior prediction (the participation decision) and post evaluation (the re-use decision) of exchange outcomes.

For the general case, much remains to be done. Having here only derived the mean exchange rate, the question of higher moments remains open. For example, computation of the variance, useful in identifying outlying events, requires an appropriate method to compute the covariances of the common name variables $c_{i_1 \cdots i_t}$ defined in section 1.1. As was seen with the case $L = 3$, this seems not to be a simple task if first finding the joint distributions is required. This hints that perhaps an alternative approach to the derivation of the second and higher moments is required.

A second major open problem is the dropping of the unbiased assumption itself. The uniform prior distribution on both appearance of resource names in state lists and the

require/release intention for each resource is the 'easy' case to consider and may rarely hold in reality. A model therefore needs to be developed for given (non-uniform) discrete distributions for these quantities. While in this case, simulation provides an alternative, the dimensionality of the parameter space describing L agents is likely to be too large to gain a simple picture of the relevant moments of the exchange variables as has been gained here in the unbiased case. An analytical result with simple reductions, as was done in section 4.5, would thus be more useful.

A number of combinatorial identities (cf. sections 3.2.2 and A.1) have resulted from the analysis of the exchange model - particularly from the case $L = 3$. It remains to be seen if these are in fact new or can be recast as known identities, perhaps as generalised convolutions, and therefore fit into an existing framework [6]. This remains a topic for further investigation.

Acknowledgements

The author would like to thank Greg Robinson (of ITG Australia Inc.) for an introduction to the crossing network modelling problem in the context of a real world application. The rest, as they say, is history.

References

1. W. E. Deming and G. J. Glasser, On the problem of matching lists by samples, *J. Am. Statist. Assoc.,* **54**, 1959, 403–415.

2. L. A. Goodman, On the analysis of samples from k lists, *Ann. Math. Statist.,* **23**, 1952, 632–634.

3. N. L. Johnson, S. Kotz and A. W. Kemp, *Univariate discrete distributions*, second edition, John Wiley and Sons, New York, 1993.

4. M. Kendall and A. Stuart, *The advanced theory of statistics*, Volume 1, fourth edition, Macmillan Publishing, New York, 1977.

5. N. N. Lebedev, *Special functions and their applications*, Dover Publications, New York, 1972.

6. R. Roy, Binomial identities and hypergeometric series, *Amer. Math. Monthly*, **97**, 1987, 36–46.

7. H. S. Wilf, *Generatingfunctionology*, Academic Press, San Diego, 1990.

Appendix A: Some Computational Details

This appendix contains quite a bit of gory detail. It provides the more computationally intensive part of section 3.2.1 ($L = 3$) by deriving an expression for $T_{lm}(r)$ - this result appears in section A.2. The results and formulae derived here are more general than are required for this task but are of interest in their own right as giving methods of computation and identities for certain sums in terms of the DG-distribution $P_2(C; r)$ discussed in section 2.2. Once these general results are established, it is an easy matter to specialise them to recover the value of $T_{lm}(r)$.

A.1 General identities

We begin by considering sums which involve the same terms as the DG-distribution $P_2(C;r)$ but summed against powers of the index. Simple versions of these sums arise in taking the expectation $E[c_{lm}c_{lq}]$ of common name variables in section 3.2.1.

Proposition 1. *For $n_l \le N_l$, $C \le B$, $A \le B$ and $s = 0, 1, 2, \dots,$ the identities*

$$\begin{aligned}
S_s(r) &:= \sum_k k^s \frac{\binom{A}{k}\binom{N_l-A}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{r}\binom{B-k}{C-r}}{\binom{B}{C}} \\
&= \sum_{n=0}^{s} \frac{(-1)^n}{n!} \frac{[n_l]_n}{[N_l]_n} \sum_{p=0}^{s} c_{n,p}(A)\Big(-\Delta_r \cdot r\Big)^p P_2(C; n_l - n, A; N_l - n, B; r) &\text{(A1)}\\
&= \sum_{n=0}^{s} \frac{[A]_n [n_l]_n}{n! [N_l]_n} \sum_{q=n}^{s} \binom{s}{q} \alpha_q \Big(-\Delta_r \cdot r\Big)^{s-q} P_2(C; n_l - n, A - n; N_l - n, B; r) &\text{(A2)}
\end{aligned}$$

hold, where

$$P_2(C; n_l, A; N_l, B; r) := \sum_k \frac{\binom{A}{k}\binom{N_l-A}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{r}\binom{B-k}{C-r}}{\binom{B}{C}}$$

is the DG*-distribution, Δ_r is the forward difference operator in r and $c_{n,p}(A)$ and α_q are expansion coefficients.*

Before proceeding with the proof, two points should be noted.

Remark. The degenerate case $B = A = C_{lm}$ (and $s = 1$) is the one of interest for the calculation of $T_{lm}(r)$ in section 3.2.1. When $B = A$, the DG-distribution reduces to the hypergeometric distribution:

$$P_2(C; n_l, A; N_l, A; r) = H_r(C; n_l, N_l) := \frac{\binom{C}{r}\binom{N_l-C}{n_l-r}}{\binom{N_l}{n_l}}, \qquad \text{(A3)}$$

for which the results (A1) and (A2) remain valid.

Remark. By defining the nonnegative quantities $\tilde{p}_k := \binom{A}{k}\binom{N_l-A}{n_l-k}/\binom{N_l}{n_l} \times \binom{k}{r}\binom{B-k}{C-r}/\binom{B}{C}$, so that $P_2(C;r) = \sum_k \tilde{p}_k$, the defining sum for $S_s(r)$ in the proposition can be regarded as a multiple of the s-th ordinary moment of the discrete (in k) distribution $\tilde{p}_k/P_2(C;r)$.

Proof of Proposition 1. The required results are derived by a straight forward application of the "snake oil" method [7]. Defining the generating function $S_s(z) := \sum_r S_s(r)z^r$, then since the hypergeometric terms $\binom{k}{r}\binom{B-k}{C-r}/\binom{B}{C}$ are generated by ${}_2F_1(-C,-k;-B;1-z)$,

$$\begin{aligned}
S_s(z) &= \sum_k k^s \frac{\binom{A}{k}\binom{N_l-A}{n_l-k}}{\binom{N_l}{n_l}} \sum_t \frac{(-C)_t(-k)_t}{(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_t \sum_k k^s [k]_t \frac{\binom{A}{k}\binom{N_l-A}{n_l-k}}{\binom{N_l}{n_l}} \frac{(-C)_t(-1)^t}{(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_t \sum_{n=0}^{s} a_{n,s}(t) \sum_k [k]_{t+n} \frac{\binom{A}{k}\binom{N_l-a}{n_l-k}}{\binom{N_l}{n_l}} \frac{(-C)_t(-1)^t}{(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_t \sum_{n=0}^{s} a_{n,s}(t)[A]_{t+n} \frac{[n_l]_{t+n}}{[N_l]_{t+n}} \frac{(-C)_t(-1)^t}{(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_{n=0}^{s} \frac{[n_l]_n}{[N_l]_n} \sum_t a_{n,s}(t)[A-t]_n \frac{(-C)_t[n_l-n]_t(-A)_t}{[N_l-n]_t(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_{n=0}^{s} \frac{(-1)^n[n_l]_n}{n![N_l]_n} \sum_t \sum_{p=0}^{s} c_{n,p}(A)t^p \frac{(-C)_t[n_l-n]_t(-A)_t}{[N_l-n]_t(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_{n=0}^{s} \frac{(-1)^n[n_l]_n}{n![N_l]_n} \sum_{p=0}^{s} c_{n,p}(A)\Big((z-1)\frac{d}{dz}\Big)^p \sum_t \frac{(-C)_t(n-n_l)_t(-A)_t}{(n-N_l)_t(-B)_t}\frac{(1-z)^t}{t!} \\
&= \sum_{n=0}^{s} \frac{(-1)^n[n_l]_n}{n![N_l]_n} \sum_{p=0}^{s} c_{n,p}(A)\Big((z-1)\frac{d}{dz}\Big)^p {}_3F_2(-C;n-n_l,-A;n-N_l,-B;1-z).
\end{aligned}$$

Since the operator $(z-1)d/dz$ on generating functions corresponds to the operator $-\Delta_r \cdot r$ on coefficients, and ${}_3F_2(-C,-m,-n;-M,-N;1-z) = \sum_r P_2(C;m,n;M,N;r)z^r$, the last identity produces the required result. A similar chain of identities but with the split $[A]_{t+n} = [A-n]_t[A]_n$ produces the equality

$$S_s(z) = \sum_{p=0}^{s} \frac{[A]_p[n_l]_p}{p![N_l]_p} \sum_{q=p}^{s} \binom{s}{q}\alpha_q \Big((z-1)\frac{d}{dz}\Big)^{s-q} {}_3F_2(-C,-n_l+p,-A+p;-B,-N_l+p;1-z)$$

which is the second form of the sum in (A2).

To make use of either of the forms (A1) or (A2) requires expressions for the coefficients $c_{n,p}(A)$ and α_q. The terms $a_{n,s}(t)$ are defined through $k^s[k]_t = \sum_{n=0}^{s} a_{n,s}(t)[k]_{t+n}$, while the terms $(-1)^n c_{n,p}(A)/n!$ are the coefficients of t^p in $a_{n,s}(t)[A-t]_n$. The degree $s-n$ polynomials $a_{n,s}(t)$ satisfying the recursion

$$\begin{aligned}
a_{0,s}(t) &= t\, a_{0,s-1}(t), \quad n = 0, \\
a_{n,s}(t) &= a_{n-1,s-1}(t+1) + t\, a_{n,s-1}(t), \quad n > 0.
\end{aligned}$$

Consequently, $a_{0,s}(t) = t^s$, $a_{1,s}(t) = (1+t)^s - t^s$, $a_{2,s}(t) = \sum_{q=0}^{s-2} t^q[(2+t)^{s-q-1} - (1+t)^{s-q-1}]$ and generally,

$$a_{n,s}(t) = \sum_{q=0}^{s-n} t^q a_{n-1,s-1-q}(1+t),$$

from which there follows the closed form expression

$$\begin{aligned} a_{n,s}(t) &= \frac{1}{n!}\sum_{q=0}^{n}(-1)^{n-q}\binom{n}{q}(t+q)^s = \frac{1}{n!}\sum_{q=n}^{s}\alpha_q\binom{s}{q}t^{s-q}, \quad \text{where} \\ \alpha_q &:= \left(x\frac{d}{dx}\right)_1^q (x-1)^n = \left(\frac{d}{dt}\right)_0^q (e^t-1)^n. \end{aligned}$$

Here, the introduction of the coefficients α_q is achieved by expanding the $(t+q)^s$ term. To finally obtain $c_{n,p}$, use the explicit form of $a_{n,s}(t)$ to expand $a_{n,s}(t)[A-t]_n$ to get

$$\begin{aligned} c_{n,p} &= \sum_{m=0}^{p}\beta_m\gamma_{p-m}, \quad p = 0,\ldots,s, \\ \beta_m &= \sum_{l=m}^{n}\begin{bmatrix} n \\ l \end{bmatrix}(-A)^{l-m}\binom{l}{m}, \quad m = 0,\ldots,n, \\ \gamma_q &= \alpha_{s-q}\binom{s}{q}, \quad q = 0,\ldots,s-n, \end{aligned}$$

with α_q as above and $\begin{bmatrix} n \\ l \end{bmatrix}$ the first kind Stirling number [7]. Observe that the two forms (A1) and (A2) originate from the alternative splits $[A]_{t+n} = [A]_n[A-n]_t = [A]_t[A-t]_n$. □

For reference, we compute and list a few of the coefficients $c_{n,p}(A)$ as $(s+1)\times(s+1)$ matrices with row index n and column index p, both beginning at zero and ending at s. Of course, for $s=0$, the only nonzero coefficient is $c_{0,0}(A) = 1$. Below are the matrices of coefficients for $s = 1,\ldots,4$:

$$s=1: \quad \begin{pmatrix} 0 & 1 \\ -A & 1 \end{pmatrix}$$

$$s=2: \quad \frac{1}{2!}\begin{pmatrix} 0 & 0 & 1 \\ -A & -(2A-1) & 2 \\ 2A(A-1) & -2(2A-1) & 2 \end{pmatrix}$$

$$s=3: \quad \frac{1}{3!}\begin{pmatrix} 0 & 0 & 0 & 1 \\ -A & -(3A-1) & -3(A-1) & 3 \\ 6A(A-1) & 6\left(A^2-3A+1\right) & -12A & 6 \\ -6A(A-1)(A-2) & 6\left(3A^2-6A+2\right) & -18(A-1) & 6 \end{pmatrix}$$

$$s=4: \quad \frac{1}{4!}\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ -A & -(4A-1) & -2(3A-2) & -2(2A-3) & 2^2 \\ 14A(A-1) & 2\left(12A^2-26A+7\right) & 2\left(6A^2-30A+19\right) & -12(2A-3) & 2^2 3 \\ -36A(A-1)(A-2) & -12(2A-3)\left(A^2-6A+2\right) & 12\left(6A^2-21A+13\right) & -36(2A-3) & 2^3 3 \\ 24[A]_4 & -48(2A-3)\left(A^2-3A+1\right) & 24\left(6A^2-18A+11\right) & -48(2A-3) & 2^3 3 \end{pmatrix}.$$

Whereas $S_s(r)$ gives the ordinary moments of $\tilde{p}_k/P_2(\mathcal{C};r)$ (see the Remark above), the factorial moments can also be considered by defining

$$C_s(r) := \sum_k \binom{k}{s} \frac{\binom{A}{k}\binom{N_l-A}{n_l-k}}{\binom{N_l}{n_l}} \frac{\binom{k}{r}\binom{B-k}{\mathcal{C}-r}}{\binom{B}{\mathcal{C}}}.$$

The defining relations

$$[k]_s = \sum_{i=0}^{s} \begin{bmatrix} s \\ i \end{bmatrix} k^i \qquad \text{and} \qquad k^s = \sum_{i=0}^{s} \begin{Bmatrix} s \\ i \end{Bmatrix} [k]_i$$

for the first and second kind Stirling numbers [7] imply the corresponding identities

$$C_s(r) = \frac{1}{s!} \sum_{i=0}^{s} \begin{bmatrix} s \\ i \end{bmatrix} S_i(r) \qquad \text{and} \qquad S_s(r) = \sum_{i=0}^{s} i! \begin{Bmatrix} s \\ i \end{Bmatrix} C_i(r).$$

These in turn lead to a companion result to (A1); namely,

$$C_s(r) = \sum_{n=0}^{s} \frac{(-1)^n}{n!} \frac{[n_l]_n}{[N_l]_n} \sum_{p=0}^{s} \tilde{c}_{n,p}(A) \Big(-\Delta_r \cdot r\Big)^p P_2(\mathcal{C}; n_l - n, A; N_l - n, B; r),$$

where the coefficients $(-1)^n \tilde{c}_{n,p}(A)/n!$ are now the coefficients of t^p in $\tilde{a}_{n,s}(t)[A-t]_n$ with

$$\tilde{a}_{n,s}(t) := \frac{1}{s!} \sum_{r=n}^{s} \begin{bmatrix} s \\ r \end{bmatrix} a_{n,r}(t) = \frac{1}{s!n!} \sum_{q=n}^{s} \frac{\alpha_q}{q!} \left(\frac{d}{dt}\right)^q [t]_s$$

and the same $a_{n,r}(t)$ as above. The first few sets of coefficients $\tilde{c}_{n,p}(A)$ are, for $s = 1, \ldots, 4$:

$$s=1: \quad \begin{pmatrix} 0 & -1 \\ -A & 1 \end{pmatrix}$$

$$s=2: \quad \frac{1}{2!} \begin{pmatrix} 0 & -1 & 1 \\ 0 & -2\,A & 2 \\ 2\,(A-1)\,A & -2\,(2\,A-1) & 2 \end{pmatrix}$$

$$s=3: \quad \frac{1}{3!} \begin{pmatrix} 0 & 2 & -3 & 1 \\ 0 & 3\,A & -3\,(A+1) & 3 \\ 0 & 6\,(A-1)\,A & -6\,(2\,A-1) & 2\cdot 3 \\ -6\,(A-2)\,(A-1)\,A & 6\,\left(3\,A^2-6\,A+2\right) & -18\,(A-1) & 2\cdot 3 \end{pmatrix}$$

$$s=4: \quad \frac{1}{4!} \begin{pmatrix} 0 & -6 & 11 & -6 & 1 \\ 0 & -8\,A & 4\,(3\,A+2) & -4\,(A+3) & 2^2 \\ 0 & -12\,(A-1)\,A & 12\left(A^2+A-1\right) & -24\,A & 2^2\,3 \\ 0 & -24\,(A-2)\,(A-1)\,A & 24\left(3\,A^2-6\,A+2\right) & -72\,(A-1) & 2^3\,3 \\ 24\,[A]_4 & -48\,(2\,A-3)\left(A^2-3\,A+1\right) & 24\left(6\,A^2-18\,A+11\right) & -48\,(2\,A-3) & 2^3\,3 \end{pmatrix}.$$

A.1.1 Recursive evaluation of (A1)

One alternative for the evaluation of the RHS of (A1) is to use the identity (for integer $n \geq 0$)

$$H_r(A, n_l - n, N_l - n) = \frac{(N_l)_n (n_l - r)_n}{(n_l)_n (N_l - A)_n} H_r(A, n_l, N_l) \tag{A4}$$

directly in the hypergeometric term in $P_2(C; n_l - n, A; N_l - n, B; r)$. By then using the expansion $[k-l]_n = (-1)^n \sum_{i=0}^{n} (-1)^i \sigma_{n-i} k^i$, where $\sigma_i := \sigma_i(l, l+1, \cdots, l+n-1)$, $0 \le i \le n$, are the elementary symmetric functions in $l, l+1, \cdots, l+n-1$, produces the equality

$$P_2(C; n_l - n, A; N_l - n, B; r) = \frac{(N_l)_n}{(N_l - A)_n} P_2(C; n_l, A; N_l, B; r) + \frac{(N_l)_n}{(n_l)_n (N_l - A)_n} \sum_{i=1}^{n} (-1)^i \sigma_{n-i} S_i(r). \tag{A5}$$

When used in (A1), (A5) expresses $S_s(r)$ in terms of $S_i(r)$, $1 \le i \le s$, and so leads to a five step procedure for the recursive evaluation of $S_s(r)$:

1. Use (A5) to write $P_2(C; n_l - n, A; N_l - n, B; r)$ in terms of $P_2(C; n_l, A; N_l, B; r)$ and $S_i(r)$, $1 \le i \le s$;
2. Substitute for $P_2(C; n_l - n, A; N_l - n, B; r)$ in (A1) of the Proposition to obtain an order s difference equation for $S_s(r)$ involving $P_2(C; n_l, A; N_l, B; r)$ and $S_i(r)$, $1 \le i \le s-1$;
3. Using the correspondence $(1-z)d/dz \leftrightarrow \Delta_r \cdot r$ between generating functions and coefficients, convert the difference equation to an order s ODE for the generating function $S_s(z)$;
4. Solve this ODE for $S_s(z)$ modulo $\mathcal{L}({}_3F_2(1-z)) = 0$, where $\mathcal{L}$ is the hypergeometric ODO for the generating function of $P_2(C; n_l, A; N_l, B; r)$;
5. Pass from the generating function $S_s(z)$ to obtain the coefficients $S_s(r)$ in terms of the $S_j(r)$, $1 \le j \le s-1$, and $P_2(C; n_l, A; N_l, B; r)$.

To illustrate, these steps are implemented for $s = 1$. Performing steps 1 and 2 gives the first order difference equation

$$\Delta_r\{r S_1(r)\} + N_l S_1(r) = (A + n_l - N_l)\Delta_r\{r P_2(r)\} + n_l A P_2(r),$$

where the abbreviation $P_2(r) \equiv P_2(C; n_l, A; N_l, B; r)$ has been made. Step 3 then gives the first order ODE

$$(1-z)S_1'(z) + N_l S_1(z) = (A + n_l - N_l)(1-z)P_2'(z) + n_l A P_2(z).$$

Applying an integrating factor, this is the same as

$$\{(1-z)^{-N_l} S_1(z)\}' = \{(1-z)^{-N_l}\left[\alpha P_2(z) + (\beta_0 + \beta_1 z)P_2'(z) + \gamma z(1-z)P_2''(z)\right]\}' - \gamma(1-z)^{-N_l-1}\mathcal{L}(P_2(z)),$$

where

$$\alpha = \frac{n_l A}{N_l - C}, \qquad \gamma = \frac{-1}{N_l - C},$$
$$\beta_0 = -1 + \frac{A + n_l - B - 1}{N_l - C}, \qquad \beta_1 = 1 + \frac{1 - A - n_l}{N_l - C},$$

and $\mathcal{L}$ is the ordinary differential operator given by

$$\mathcal{L}(u) = z(1-z)^2u''' - (1-z)[1 - N_l - B - (3 - \mathcal{C} - n_l - A)(1-z)]u'' + [N_lB - (1 - \mathcal{C} - n_l - A + n_l\mathcal{C} + n_lA + A\mathcal{C})(1-z)]u' - n_l\mathcal{C}Au.$$

Since the generating function $P_2(z) \equiv {}_3F_2(-\mathcal{C}; -n_l, -A; -N_l, -B; 1-z)$ lies in the kernel of $\mathcal{L}$ [5], then integrating gives the relation (step 4)

$$S_1(z) = \alpha P_2(z) + (\beta_0 + \beta_1 z)P_2'(z) + \gamma z(1-z)P_2''(z).$$

Finally, returning to the generating function's coefficients (step 5) gives the relation

$$S_1(r) = \big[r + \frac{(n_l - r)(A - r)}{N_l - \mathcal{C}}\big]P_2(r) - (r+1)\big[1 + \frac{1 - A - n_l + B + r}{N_l - \mathcal{C}}\big]P_2(r+1). \quad \text{(A6)}$$

This expresses $S_1(r)$ as a simple quadratic (in r) combination of $P_2(\mathcal{C}; r)$ with index shifts of 0 and 1.

A.2 Evaluation of $T_{lm}(r)$

This section gives two derivations of the value of $T_{lm}(r)$ corresponding to the two techniques above for the evaluation of $S_1(r)$. The case of interest is $B = A$ and hence, as noted in (A3), the reduction $P_2(\mathcal{C}; A, n_l; N_l, A; r) = H_r(\mathcal{C}, n_l, N_l)$ holds.

A.2.1 Method 1

For $s = 1$, the coefficients $c_{0,0} = 0$, $c_{0,1} = 1$, $c_{1,0} = -A$ and $c_{1,1} = 1$ from above can be substituted into (A1) to give

$$S_1(r) = -\Delta_r \cdot rH_r(\mathcal{C}, n_l, N_l) + (-1)\frac{n_l}{N_l}\Big\{-A + (-\Delta_r \cdot r)\Big\}H_r(\mathcal{C}, n_l - 1, N_l - 1). \quad \text{(A7)}$$

Now the hypergeometric probabilities H_r satisfy (put $n = 1$ in (A4))

$$\frac{n_l}{N_l}H_r(\mathcal{C}, n_l - 1, N_l - 1) = \frac{n_l - r}{N_l - \mathcal{C}}H_r(\mathcal{C}, n_l, N_l)$$

and

$$-\Delta_r rH_r(\mathcal{C}, n_l - k, N_l - k) = \frac{(N_l - k + 1)r - (n_l - k)\mathcal{C}}{N_l - \mathcal{C} + r - n_l + 1}H_r(\mathcal{C}, n_l - k, N_l - k),$$

for $k = 0, 1, \ldots$, which, with $k = 1$ simplifies (A7) to

$$S_1(r) = \Big\{A\frac{n_l - r}{N_l - \mathcal{C}} + \frac{rN_l - n_l\mathcal{C}}{N_l - \mathcal{C}}\Big\}H_r(\mathcal{C}, n_l, N_l).$$

Setting $A = \mathcal{C}_{lm}$ finally gives the required result

$$T_{lm}(r) = \frac{\mathcal{C}_{lm}(n_l - r) + rN_l - n_l\mathcal{C}}{N_l - \mathcal{C}} \; \frac{\binom{\mathcal{C}}{r}\binom{N_l - \mathcal{C}}{n_l - r}}{\binom{N_l}{n_l}}.$$

A.2.2 Method 2

From the expression (A6) for $S_1(r)$ derived by the ODE method, we need only set $B = A = C_{lm}$ and substitute H_r for $P_2(r)$. Since H_r satisfies the recursion

$$H_{r+1}(\mathcal{C}, n_l, N_l) = \frac{(\mathcal{C} - r)(n_l - r)}{(r+1)(N_l - \mathcal{C} - n_l + r + 1)} H_r(\mathcal{C}, n_l, N_l),$$

then substitution into the RHS of (A6) produces the same result for $T_{lm}(r)$.

Appendix B: Derivations and Identities

After deriving the three-agent common name distribution in section B.1, section B.2 of this appendix lists some elementary but useful combinatorial identities, including what is called the Combinatorial Lemma (CL). Finally, in section B.3, an alternative derivation of the mean exchange rate for all agents is given.

B.1 Derivation of $P(c_{lm}, c_{lq})$ for $L = 3$

In this section the LTP and Bayes' rule are applied to derive the joint distribution $P(c_{lm}, c_{lq})$ of common names as stated in section 3.2.1. Define the notation $P(k \text{ of } C_{lm} \text{ in } l)$ to be the probability that k of the C_{lm} names common to pools P_l and P_m are in the state list of agent a_l, etc. By the LTP,

$$P(c_{lm}|c_{lq}) = \sum_k P(c_{lm}|k \text{ of } C_{lm} \text{ in } l)P(k \text{ of } C_{lm} \text{ in } l|c_{lq}),$$

where the first probability in this sum is $\binom{k}{c_{lm}}\binom{N_m-k}{n_m-c_{lm}}/\binom{N_m}{n_m}$. Considering the second term in the sum and applying Bayes' rule gives

$$\begin{aligned}&P(k \text{ of } C_{lm} \text{ in } l|c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q) = \\ &\quad P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q|k \text{ of } C_{lm} \text{ in } l) \times \frac{P(k \text{ of } C_{lm} \text{ in } l)}{P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q)}\end{aligned} \tag{B1}$$

where the ratio of the two terms on the RHS is just:

$$\frac{P(k \text{ of } C_{lm} \text{ in } l)}{P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q)} = \frac{\binom{C_{lm}}{k}\binom{N_l-C_{lm}}{n_l-k}}{\binom{N_l}{n_l}P_2(C_{lq}; c_{lq})}.$$

Now consider the first term on the RHS of (B1) and again apply the LTP (twice) to get

$$\begin{aligned}&P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q|k \text{ of } C_{lm} \text{ in } l) = \\ &\sum_r P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q|r \text{ of } C \text{ in } l)P(r \text{ of } C \text{ in } l|k \text{ of } C_{lm} \text{ in } l) \\ &= \sum_{r,s} P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q|s \text{ of } C_{lq} \text{ in } l)P(s \text{ of } C_{lq} \text{ in } l|r \text{ of } C \text{ in } l) \\ &\qquad \times P(r \text{ of } C \text{ in } l|k \text{ of } C_{lm} \text{ in } l).\end{aligned}$$

The first term on the RHS is $\binom{s}{c_{lq}}\binom{N_q-s}{n_q-c_{lq}}/\binom{N_q}{n_q}$ while applying Bayes' rule to the second gives

$$P(s \text{ of } C_{lq} \text{ in } l|r \text{ of } C \text{ in } l) = \frac{P(s \text{ of } C_{lq} \text{ in } l)}{P(r \text{ of } C \text{ in } l)}P(r \text{ of } C \text{ in } l|s \text{ of } C_{lq} \text{ in } l). \tag{B2}$$

Now the numerator of the ratio in the RHS of (B2) is $\binom{C_{lq}}{s}\binom{N_l-C_{lq}}{n_l-s}/\binom{N_l}{n_l}$ while the denominator is $\binom{C}{r}\binom{N_l-C}{n_l-r}/\binom{N_l}{n_l}$. The remaining terms in the product are simply

$$P(r \text{ of } C \text{ in } l|s \text{ of } C_{lq} \text{ in } l) = \frac{\binom{s}{r}\binom{C_{lq}-s}{C-r}}{\binom{C_{lq}}{C}}$$

and

$$P(r \text{ of } C \text{ in } l|k \text{ of } C_{lm} \text{ in } l) = \frac{\binom{k}{r}\binom{C_{lm}-k}{C-r}}{\binom{C_{lm}}{C}}.$$

Finally, putting all of these terms together and multiplying through by the probability density $P(c_{lq} \text{ of } C_{lq} \text{ in } l \text{ and } q) \equiv P_2(C_{lq}; c_{lq})$ gives the desired expression for $P(c_{lm}, c_{lq})$ in (5).

B.2 Elementary identities

The derivation of the means in section 4 relies on simple identities involving multi-indexed summations. Although only ever using part (a) in section 4, the results below are listed for reference. Parts (b) and (c) have applications in the following section and in reductions of the means as in section 4.5.

Combinatorial Lemma (CL). *For any multi-indexed quantities $Q_{i_1\cdots i_t}$, $2 \le t \le L$, the following identities hold:*

(a) For $0 \le r + k \le L$,

$$\sum_{i_1\ldots i_k} \sum_{\substack{j_1\ldots j_r \not\ni \\ i_1,\cdots,i_k}} Q_{i_1\ldots i_k j_1\ldots j_r} = \binom{r+k}{k} \sum_{i_1\ldots i_{k+r}} Q_{i_1\ldots i_{k+r}};$$

(b) If $s(\lambda_{[k]}, r)$ is the r-th symmetric polynomial in variables $\lambda_{i_1}, \cdots, \lambda_{i_k}$, then for $r \le k \le L$ and $0 \le t \le L - k$,

$$\sum_{i_1\cdots i_k} \sum_{\substack{j_1\cdots j_t \not\ni \\ i_1,\cdots,i_k}} Q_{i_1\cdots i_k j_1\cdots j_t}\, s(\lambda_{[k]}, r) = \binom{k+t-r}{t} \sum_{i_1\cdots i_{k+t}} Q_{i_1\cdots i_{k+t}} s(\lambda_{[k+t]}, r);$$

(c) If $s(\lambda_{[k]}, t)$ is the t-th symmetric polynomial in variables $\lambda_{i_1}, \cdots, \lambda_{i_k}$, then

$$\sum_{\substack{i_1,\cdots,i_k \in \\ \{1,\cdots,L\}}} s(\lambda_{[k]}, t) = \binom{L-t}{k-t} s(\lambda_{[L]}, t), \quad t = 0, \cdots, k.$$

The proofs of these facts follow by simple counting arguments on the possible number of combinations. The following elementary result is also noted.

Lemma. *For the finite sequence* $p_0, \dots, p_n$*, then the identities*

$$b_k := \sum_{r=0}^{n} \binom{r}{k} p_r$$
$$p_r = \sum_{k=r}^{n} (-1)^{k-r} \binom{k}{r} b_k$$

hold. Additionally, for any positive integer s *and with* b_k *defined as above, then*

$$k^s b_k = \sum_{r=0}^{n} \binom{r}{k} (-\Delta_r \cdot r)^s p_r$$
$$(-\Delta_r \cdot r)^s p_r = \sum_{k=r}^{n} (-1)^{k-r} \binom{k}{r} k^s b_k$$

where Δ_r *is the forward difference operator.*

This Lemma follows from an application of the 'snake oil' method [7]. When the sequence $\{p_r\}$ is a probability density, then $k!b_k = \mu'_{[k]}$ are the factorial moments and hence p_r is written in terms of these moments. This gives another easy way of recognising the factorial moments of the DG-distribution from the expression for $P_t(\mathcal{C}; c)$ in section 2.2.

B.3 Alternative derivation for all agents

An alternative derivation for the expression for μ in section 4.1 which is based upon a consideration of the overlaps of resource pools rather than state lists is detailed here. Although this approach is a little more involved, it is included as a matter of side interest since, historically, it provided the first derivation. Some slight notational differences to the main body of the paper are used.

Let $j_{i_1 \cdots i_t}$ denote the number of exchangeable names in the state lists of agents $a_{i_1}, \dots, a_{i_t}$ among those names uniquely common to the *pools* $P_{i_1}, \dots, P_{i_t}$. Then the total number of exchangeable names is

$$j = \sum_{t=2}^{L} \sum_{i_1 \cdots i_t} j_{i_1 \cdots i_t}. \tag{B3}$$

The idea here then is to add up the number of exchangeable names in the uniquely common names to all pools rather than state lists. By an argument analogous to the derivation of (12), the number $C^U_{i_1 \cdots i_k}$ of names uniquely common to pools $P_{i_1}, \dots, P_{i_k}$ is just

$$C^U_{i_1 \cdots i_k} = \sum_{t=0}^{k} (-1)^t \sum_{\substack{j_1 \cdots j_t \not\ni \\ i_1, \dots, i_k}} C_{i_1 \cdots i_k j_1 \cdots j_t} \tag{B4}$$

which also gives the upper bound on $j_{i_1\cdots i_k}$. From this expression, for each fixed $\ell = 1,\ldots,L$, if C_ℓ^U denotes the number of names in pool P_ℓ common to all other pools, then since the total overlap of pool P_ℓ with all other pools cannot exceed its size, the inequalities

$$C_\ell^U := \sum_{k=1}^{L-1}\sum_{i_1\cdots i_k} C^U_{\ell i_i\cdots i_k} = \sum_{k=1}^{L-1}(-1)^{k+1}\sum_{i_1\cdots i_k} C_{\ell i_i\cdots i_k} \le N_\ell$$

must hold for each $\ell = 1,\ldots,L$. This provides a consistency check on the specification of the overlap parameters $C_{i_1\cdots i_k}$.

To derive $\mu = E[j]$, first requires $E[j_{i_1\cdots i_t}]$. To this end, we first show that

$$E[j_{i_1\cdots i_k}] = C^U_{i_1\cdots i_k}\sum_{t=2}^{k}(-1)^t(1-2^{1-t})s(\lambda_{[k]},t) \tag{B5}$$

where $s(\lambda_{[k]},t)$ is the t-th symmetric polynomial in $\lambda_{i_1},\ldots,\lambda_{i_k}$. To minimise the subscript notation, it suffices to prove this for $k = L$ (which can be replaced by any smaller subset of the L agents). Among the $C^U_{1\cdots L} \equiv C$ names uniquely common to pools $P_1,\ldots,P_L$, let $c'_{i_1\cdots i_k}$ be the number of these names common to, and $d'_{i_1\cdots i_k}$ the number of these names uniquely common to, the state lists of agents $a_{i_1},\ldots,a_{i_k}$, and let $j'_{i_1\cdots i_k}$ be the number of exchangeable names among the $d'_{i_1\cdots i_k}$ uniquely common ones. Then by (ii) of the PL, $E[j'_{i_1\cdots i_k}] = (1-2^{1-k})E[d'_{i_1\cdots i_k}]$. Now, since

$$d'_{i_1\cdots i_k} = c'_{i_1\cdots i_k} - \sum_{t=1}^{L-k}\sum_{\substack{j_1\cdots j_t \not\ni \\ i_1,\ldots,i_k}} d'_{i_1\cdots i_k j_1\cdots j_t}$$

then, through an argument the same as in section 4, and using (c) of the CL, $\chi'_k := \sum_{i_1\cdots i_k} E[d'_{i_1\cdots i_k}]$ is given by

$$\chi'_k = C\sum_{r=k}^{L}(-1)^{r-k}\binom{r}{k}s(\lambda_{[L]},r).$$

Now, writing the decomposition

$$j_{1\cdots L} = \sum_{k=2}^{L}\sum_{i_1\cdots i_k} j'_{i_1\cdots i_k}$$

gives $E[j_{1\cdots L}] = \sum_{k=2}^{L}(1-2^{1-k})\chi'_k$. Substitution for χ'_k from above, and switching the summations, then gives the required result in (B5) for $k = L$.

Substitution of (B4) into (B5) and the result into the expectation of (B3) gives

$$\begin{aligned}\mu &= \sum_{k=2}^{L}\sum_{t=2}^{k}(-1)^t(1-2^{1-t})\sum_{r=0}^{L-k}(-1)^r\sum_{i_1\cdots i_k}\sum_{\substack{j_1\cdots j_r \not\ni \\ i_1,\ldots,i_k}} C_{i_1\cdots i_k j_1\cdots j_r}s(\lambda_{[k]},t)\\ &= \sum_{t=2}^{L}(-1)^t(1-2^{1-t})T(t)\end{aligned}$$

where, having used (b) of the CL,

$$T(t) = \sum_{r=t}^{L}(-1)^r \sum_{i_1\cdots i_r} C_{i_1\cdots i_r} s(\lambda_{[r]}, t) \sum_{k=t}^{r}(-1)^k \binom{r-t}{r-k}.$$

Now, since $\sum_{k=t}^{r}(-1)^k\binom{r-t}{r-k} = (-1)^t\delta_{rt}$, only the $r = t$ term survives in $T(t)$ and so

$$T(t) = \sum_{i_1\cdots i_t} C_{i_1\cdots i_t}\lambda_{i_1}\cdots\lambda_{i_t}$$

where the substitution $s(\lambda_{[t]}, t) = \lambda_{i_1}\cdots\lambda_{i_t}$ has been made. This then reproduces the expression (9) from section 4.1.

DISTRIBUTION LIST

Combinatorics of Crossing Networks

Geoff A. Latham

	Number of Copies
DEFENCE ORGANISATION	
S&T Program	
Chief Defence Scientist FAS Science Policy AS Science Corporate Management	1
Director General Science Policy Development	1
Counsellor, Defence Science, London	Doc Data Sht
Counsellor, Defence Science, Washington	Doc Data Sht
Scientific Adviser to MRDC, Thailand	Doc Data Sht
Scientific Adviser Policy and Command	1
Navy Scientific Adviser	Doc Data Sht
Scientific Adviser, Army	Doc Data Sht
Air Force Scientific Adviser	1
Director Trials	1
Aeronautical and Maritime Research Laboratory	
Director, Aeronautical and Maritime Research Laboratory	1
Electronics and Surveillance Research Laboratory	
Director, Electronics and Surveillance Research Laboratory	Doc Data Sht
Chief, Communications Division	1
Chief, Information Technology Division	1
Research Leader, Military Information Networks, CD	1
Head, Network Requirements, CD	1
Research Leader, Secure Communications, CD	1
Head, Intelligent Networks, CD	1
Head, Information Access, CD	1
Head, C3I Networks, CD	1
Head, Crypto-Mathematics Research, CD	1
Head, Information Management and Fusion, ITD	1
Mr. John Ross, CD (Task Manager)	1
Dr. Geoff Latham, CD (Author)	5
Dr. Martin Oxenham, SSD	1
DSTO Research Library and Archives	
Library Fishermans Bend	1

Library Maribyrnong	1
Library Salisbury	2
Australian Archives	1
Library, MOD, Pyrmont	Doc Data Sht
US Defense Technical Information Center	2
UK Defence Research Information Centre	2
Canada Defence Scientific Information Service	1
NZ Defence Information Centre	1
National Library of Australia	1
Capability Systems Staff	
Director General Maritime Development	Doc Data Sht
Director General Land Development	Doc Data Sht
Director General C3I Development	Doc Data Sht
Director General Aerospace Development	Doc Data Sht
Navy	
SO(Science), Director of Naval Warfare, Maritime Headquarters Annex, Garden Island	Doc Data Sht
Army	
ASNSO ABCA, Puckapunyal	4
SO(Science), DJFHQ(L), MILPO, Enoggera, Queensland 4057	Doc Data Sht
Air Force	
Intelligence Program	
DGSCD, Defence Signals Directorate	1
DGSTA, Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1
Acquisitions Program	
Corporate Support Program	
Officer in Charge, TRS, Defence Regional Library, Canberra	1
UNIVERSITIES AND COLLEGES	
Australian Defence Force Academy Library	1
Head of Aerospace and Mechanical Engineering, ADFA	1
Deakin University Library, Serials Section (M List)	1
Hargrave Library, Monash University	Doc Data Sht
Librarian, Flinders University	1

OTHER ORGANISATIONS

NASA (Canberra)	1
Info Australia	1
State Library of South Australia	1
Parliamentary Library of South Australia	1
CSIRO, Mathematics and Information Sciences, North Ryde, Sydney	1
Investment Technology Group, Australia Inc., Melbourne	1

ABSTRACTING AND INFORMATION ORGANISATIONS

Library, Chemical Abstracts Reference Service	1
Engineering Societies Library, US	1
Materials Information, Cambridge Scientific Abstracts, US	1
Documents Librarian, The Center for Research Libraries, US	1

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK	1
Library – Exchange Desk, National Institute of Standards and Technology, US	1
National Aerospace Laboratory, Japan	1
National Aerospace Laboratory, Netherlands	1

SPARES

DSTO Salisbury Research Library	5

Total number of copies: 66

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA					1. CAVEAT/PRIVACY MARKING

2. TITLE	3. SECURITY CLASSIFICATION
Combinatorics of Crossing Networks	Document (U) Title (U) Abstract (U)
4. AUTHOR(S) Geoff A. Latham	5. CORPORATE AUTHOR Electronics and Surveillance Research Laboratory PO Box 1500 Salisbury, South Australia, Australia 5108

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–RR–0176	AR-011-477	Research Report	June 2000

8. FILE NUMBER	9. TASK NUMBER	10. SPONSOR	11. No OF PAGES	12. No OF REFS
E8730/16/34	DST 98/224	CDS	45	7

13. URL on the Worldwide Web	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-RR-0176.pdf	Chief, Communications Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, SALISBURY, SOUTH AUSTRALIA 5108

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DEFTEST DESCRIPTORS

Combinatorial analysis
Exchanging
Performance prediction
Resources management

19. ABSTRACT

In scenarios such as software agents sharing time slots or blocks in different physical CPUs or memories, communications agents accessing channels or bandwidth on shared links or bearers and economic agents trading in commodities in financial markets, the agents in question can exchange resources among themselves for mutual benefit. This report develops a combinatorial probability model to describe the like-for-like exchange of resources between multiple unbiased agents. The expected exchange rates are computed for individual agents, syndicates of agents and the collection of all agents. These performance benchmarks provide intuition and understanding for the performance of real crossing networks. A number of combinatorial identities are also produced as a consequence of the modelling and analysis.

Page classification: UNCLASSIFIED